

Time for fair trading of supercomputers

The United States should follow internationally accepted procedures in dealing with the alleged 'dumping' of a Japanese supercomputer system rather than yield to the kangaroo-court tactics of congressmen with vested interests.

WOULD the lease of one Japanese supercomputer system to the US National Center for Atmospheric Research (NCAR) threaten the future economic strength and national security of the United States? Apparently so, if one is to believe the arguments of some US congressmen. Last week, they pushed forward legislation that will deny salaries to any officials of the US National Science Foundation (NSF) who approve the centre's procurement of an NEC supercomputer system — if it is subsequently established that the Japanese computers are being 'dumped' at below cost (see page 5).

Led by David Obey of Wisconsin and Martin Sabo of Minnesota, Democrats in whose districts the US supercomputer manufacturer Cray Research is based, some members of Congress claim that the threat is so great and the case against NEC so clear-cut that the United States must bypass internationally recognized procedures for dealing with dumping and nip the procurement in the bud with their custom-made amendment tagged onto the NSF appropriations bill. But their action seems to have more to do with the failing strength of Cray — which has just been taken over by Silicon Graphics of California — than with any immediate threat to the United States.

What evidence is there that the NEC supercomputer system is being dumped? So far, none that is unbiased. Opponents of the NEC purchase have made some back-of-the-envelope calculations which suggest that the NEC supercomputer system being offered to the centre for US\$35 million is worth between US\$90 and US\$110 million. And two weeks ago, the *Nihon Kogyo Shimbun*, a leading Japanese industrial newspaper, released a more detailed analysis, attributed to the US Department of Commerce, which suggests the loss for NEC is more than US\$100 million (see *Nature*, 381, 723; 1996). But the Commerce Department disowns these figures, and subsequent investigation by *Nature* has revealed that the newspaper obtained this analysis from Cray Japan.

Countering these calculations, NEC has provided figures to NCAR that suggest the computers are not being offered at below fair value. But those have not been made public.

Only a thorough investigation by an impartial body is going to establish whether the NEC supercomputers are being dumped. There are well established procedures under the World Trade Organisation (WTO) agreement to do that. But they will take time, possibly as long as a year.

Obey and his supporters claim that the United States cannot wait that long. Their amendment to the NSF appropriations bill, passed by the House of Representatives last week, is "critical to the national security interests of the United States", they say. They argue that, if one NEC supercomputer system is allowed, a flood will follow and the United States will lose control of the supercomputer industry to Japan.

A cooler analysis, however, does not support such a scenario. Cray has so far sold 320 supercomputers in the United States against NEC's two, and Japanese manufacturers are locked out of the large market for US defence-related procurements on grounds of national security. Even in Japan, which Obey and colleagues claim is still not fully open to US supercomputer manufacturers, Cray has sold more supercomputers than NEC.

But, despite profits in Japan, the mother company has been on shaky ground in recent years, with losses of more than \$200 million

last year. Silicon Graphics, however, has come to the rescue and there is no longer such a pressing need for Cray to be propped up by the US government.

A few years ago, there might have been more international sympathy for Obey and his supporters in the US Congress and Department of Commerce. In the 1980s, Japan's government procurement of supercomputers was virtually closed to US manufacturers. But, after much political arm-twisting by the United States, an agreement was reached on supercomputer trade in 1990 that has led to the purchase of dozens of Cray supercomputers by Japanese government research organizations.

As some US congressmen opposing the Obey-Sabo amendment argued last week, Congress's move to block the NEC procurement may well violate WTO agreements on dumping (a view shared by Japan's ministries of international trade and industry and foreign affairs). Furthermore, it runs completely counter to the spirit of the US-Japan agreement on supercomputer trade.

If this NSF legislation becomes law — it has yet to be passed by the Senate and will then have to be approved by the US president — Japan would have strong grounds for protesting to the WTO. The United States has been the world's strongest champion of free trade and it should not jeopardize that position for the sake of one supercomputer sale.

Regrettably, the damage has probably already been done. Obey and his colleagues have sent a powerful message to the NSF that they should buy American, and it will require considerable nerve on NSF's part to ignore that call and to stand by principles that are in the best long-term interests of US science and the US economy. □

Changing *Nature* (contd)

***Nature* is taking steps to enhance the readability of its scientific papers.**

FROM this week, *Nature* is introducing the first set of a series of changes aimed at making our scientific content more readable.

Hitherto, details of experimental methods too cumbersome to be included in the body of text of Letters and Articles have been published in figure captions. We are now phasing in the use of a separate Methods section in such papers. The section should be no longer than 800 words, positioned at the end of the text.

We have also abolished the nominal limits of numbers of display items and words, and now specify length limits in terms of *Nature* pages: normally, 2.5 pages for Letters, 5 for Articles. (A page of undiluted text in a *Nature* paper would contain 1,300 words.)

Finally, production changes have led to improvements in figure reproduction, which will be optimized if authors supply figures in digital form at the time of acceptance of their papers. Guidance on formats is available from the production department, which can be contacted via e-mail at nature@nature.com.

A correspondingly changed Guide to Authors is on page 94, and, in expanded form, on our World Wide Web sites at <http://www.nature.com> and www.america.nature.com □



Australian university fights off research council's 'hijack' bid

Sydney. A bruising battle has broken out between the Australian National University (ANU) and the Australian Research Council (ARC) over who should control the university's research finance and agenda. ANU leaders are accusing the ARC, Australia's major national grant-giving body, of trying to 'hijack' some of their funds in order to reallocate them to other universities.

Celebrating its 50th anniversary this year, the ANU is unusual in coming under the direct responsibility of the federal government. All other universities operate under the responsibility of individual state parliaments (of which there are six), although for more than two decades their funding has come from federal government agencies.

Through the research and postgraduate focus of its Institute of Advanced Studies (IAS), the ANU, located in Canberra, has long prided itself as the flagship for basic research in Australia's higher education sector. But the expansion of the university system over the past 30 years has fuelled the ambitions of other universities to become equal players in competition for research support.

About 30 per cent of operating grants to the universities — now totalling A\$4.6 billion (US\$3.6 billion) annually — was intended to support research, while a parallel scheme for allocating grants to individuals and small groups expanded through what has become the ARC. The annual competition for these latter funds, which amount to A\$350 million this year, has become intense.

In contrast, the IAS has enjoyed special treatment. This is because it is financed directly by the government through a block grant, which is distributed among eight specialized schools dedicated to research and postgraduate training, in areas ranging from Earth sciences to medical research, as well as centres for astronomy, and resource and environment studies.

Staff of the IAS, who often enjoy high academic reputations both nationally and internationally, stoutly defend the value of being able to decide their own research agenda. But the major metropolitan universities have steadily built their research bases to such an extent that they feel they are now in a position to challenge the IAS's claim to special treatment through the ARC.

In order both to plan its next seven years of activity and to develop a defence against potential challenges, the ANU set up a wide-ranging external review of the IAS in 1994, with the ARC as a participant. The review committee was chaired by Keith

Boardman, a former chairman of the Commonwealth Scientific and Industrial Organisation, and involved visits by 70 independent researchers from Australia and overseas.

In a report presented last year, the joint committee commented favourably on the quality of the institute, based on a variety of performance indicators. The review recommended the continuation of block funding, as well as a modest budget increase of A\$2 million per annum. It also said that IAS should not be disrupted with another major review for seven years.

But the ARC, chaired by Max Brennan, a plasma physicist formerly of the University of Sydney, had already got its foot in the door. At the request of the previous Labor government, the council produced a separate report, which has just been released by the new government. The ARC wants to determine the block allocation to the IAS, which came to A\$151 million in 1995, and says that this should not be increased.

At the same time, the ARC asks for its own funding from the government to be increased by A\$3.5 million year to boost its programmes for special research initiatives and infrastructure, saying that the IAS should be allowed to compete with other universities for money from these pools.

The reaction has been fierce. The ANU claims that the research council would, in effect, control the research agenda of the IAS. The university council has asked John Howard, Australia's new Prime Minister, to ignore the ARC proposals and accept those of its own joint review.

Susan Serjeantson, Director of the IAS, describes the ARC's report, to which she says the ANU was not privy, as "an audacious grab for power and an increase in [the ARC's] resources". If its proposals are accepted, she says, the research council "would control research in the higher education system almost absolutely".

The Australian Vice-Chancellors' Committee is staying outside the ANU/ARC fight, characterizing it "as purely an institutional affair". Indeed, leaders of other universities have been silent on the ANU's position, which is blurred by its membership of the newly resurgent 'Group of Eight' major metropolitan universities claiming pre-eminence in research.

The group appears to have broken ranks with the 28 other universities by meeting with Amanda Vanstone, the education minister, and publicly pressing for the diversification of higher education into predominantly research or teaching universities. It is lobbying for protection from feared cuts of up to A\$550 million in operating grants, and has proposed raising revenue through advance fees for courses, and increasing students' deferred contributions.

Brennan describes the ANU's reactions as "revealing paranoia within the institute", and denies that the ARC is just a rubber stamp. "Nothing in our report substantiates the claim of a takeover of the IAS agenda," he says. "Our report recognises the unique place of the IAS in the Australian university system, and makes constructive recommendations to enhance this role." **Peter Pockley**

Now crop circles weave a double helix

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

News Team

London. **A mysterious double-helix pattern (above), 200 metres long, appeared overnight in a field of wheat last week near the English village of Alton Barnes in Wiltshire. The pattern, which contains 89 circles, was formed in under four hours, says Polly Carson, who owns the field. "There was nothing there at midnight," she says. "The next time we looked, at 4 a.m., it was daylight, and there it was."**

Carson says crop circles regularly appear in her fields, and she does not believe they are all hoaxes, as is frequently claimed. "Some are truly amazing," she adds. "We had a scale map of the Solar System. Everything in it was accurate, even the eccentricity of the planetary orbits and their distance from the Sun." E. M.

Statistics suggest BSE now 'Europe-wide'

Paris. Europe's veterinary surgeons have urged the European Union (EU) to introduce and enforce measures to ensure that bovine spongiform encephalopathy (BSE) cannot enter the human food chain. Their call comes as new analyses suggest that the incidence of BSE in continental Europe is much higher than officially reported.

EU member states in continental Europe have declared a total of around 50 cases of BSE (see table, right). But some researchers estimate that these countries ought to have declared more than 2,000 cases, given that they imported large quantities of breeding cattle and potentially contaminated meat and bone meal from the United Kingdom (see *Nature* 381, 544; 1996).

One explanation for this discrepancy, say observers, is that farmers may have been discouraged from reporting cases because many countries slaughter herds where BSE has occurred without adequately compensating farmers.

These estimates of the number of BSE cases in continental Europe are nonetheless an order of magnitude lower than in the United Kingdom, where around 160,000 cases have been reported. It follows that any impact on human health would be correspondingly lower, even if it was established that the agent that causes BSE can pass to humans and cause Creutzfeldt-Jakob disease (CJD). Experts also point out that most of the BSE cases in continental Europe will already have entered the food chain.

Nevertheless, at a meeting in Paris last month, the Federation of Veterinarians of Europe (FEV) acknowledged the BSE crisis is now a "European problem" — and not just a British one — according to one official present. Tougher European measures to protect public health are needed, he argues.

He points out that, until last week, when France banned the human consumption of specified bovine offals (SBO) — the most infectious parts, such as brain and spinal cord — (see page 5), the United Kingdom had been the only EU country to have done so. (Switzerland, which is not a member of the EU, had introduced a similar ban.)

To reassure consumers, "it must be demonstrated beyond any doubt that BSE cannot enter the food chain", says Francis Antony, president of FEV and the British Veterinary Association's spokesman on BSE. In a statement last week, FEV called on the EU to "design, apply and enforce" a comprehensive food hygiene programme "from the stable to the table".

"Consumer perception and the scientific reality of food safety are too far apart," says one FEV official. "The key issue is about keeping brain and spinal cord out of the food chain." Indeed, the FEV statement seems a thinly veiled attack on what one official says is the EU's excessive focus on the culling of UK cattle.

Although culling will precipitate the decline of the epidemic in cattle, it will have little direct impact on human health, argues one FEV official. Eliminating risk to public health requires a ban on the human consumption of SBOs and strict enforcement of deboning and denervating procedures in abattoirs to eliminate infective tissues.

The EU also needs to ensure that adequate resources are made available to enforce any new regulations, says one FEV official. The failure of the United Kingdom to police its early feed and SBO bans has shown that "you can ban things until you are blue in the face", but without proper enforcement this has little impact.

The revised estimates of cases of BSE in mainland EU states come from analyses of UK exports of animal feed and live cattle. Experts predict that imported feed may have resulted in several hundred cases of BSE. But the total is difficult to estimate precisely because the final destination and use of feed are often unknown, while the relationship between tonnage and infectivity is difficult to establish, according to John Wilesmith, of the UK Ministry of Agriculture's Central Veterinary Laboratory in Weybridge.

Better estimates may come from analysis of UK exports of breeding cattle. One group of researchers has estimated, for example, that the United Kingdom exported 57,900 pure-bred breeding bovines between 1985

and 1990, the latter being the year after the EU banned imports of breeding animals older than six months.

One of the scientists, B. E. C. Schreuder from the DLO-Institute for Animal Science and Health in Lelystad, the Netherlands, says that, statistically, around 1,700 of these animals would be expected to have developed BSE.

Such estimates must be taken with caution because of the assumptions they involve, says Schreuder. His estimates of total exports are somewhat higher than records of exports for "pure-bred breeding

Number of declared cases of BSE worldwide (May 1996)	
UK	158,866
Switzerland	211
Eire	125
France	18
Portugal	30*
*excluding 6 cases in imported animals	

World Animal Health Organization

bovines" (see figure, below left) provided by Eurostat — the body that supplies the European Commission with statistics — as he has also included other export categories such as "bulls not intended for slaughter".

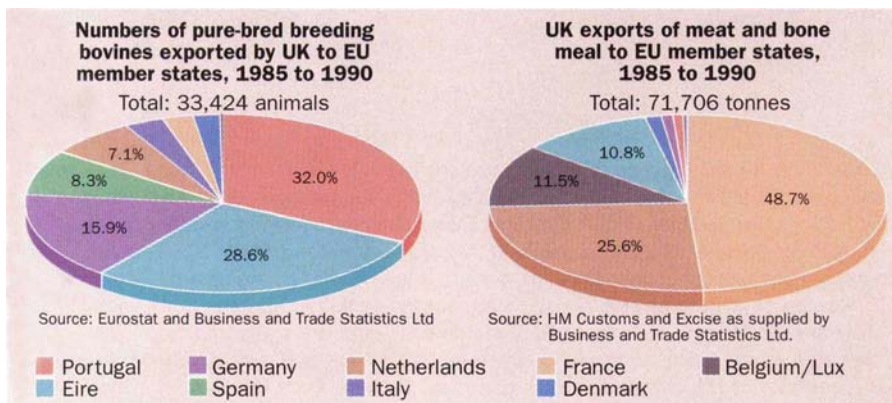
Nonetheless, Schreuder is confident that their estimate of BSE cases in continental Europe attributable to cattle exports is of the right order of magnitude. He says it is possible to identify the individual member states most at risk — Portugal, which has declared just 30 cases of BSE, is expected to top this list.

Schreuder says that his warning of under-reporting is mainly aimed at countries where lack of regulation might yet allow recycling of BSE and new outbreaks. Wilesmith adds that better data on true BSE levels in continental Europe will be essential to studying the epidemiology of any cases of the new variant of CJD that arise there. The evidence for higher levels of BSE on the European mainland also underlines the need for an effective European CJD surveillance network, he says (see *Nature* 381, 453; 1996).

Wilesmith asserts that "we were lucky there were not BSE epidemics elsewhere". The factors that led the United Kingdom to be worst affected, he says, probably included the world's highest ratio of sheep to cattle (45 million to 12 million), a high prevalence of scrapie, and the use of high proportions of meat and bone meal in cattle feed.

Other countries may have escaped relatively unscathed only because they had more cattle than sheep, low levels of scrapie, or used less meat and bone meal in cattle feed. Indeed, a new report from the UK Institute of Animal Health suggests that the procedures used in many rendering plants throughout Europe would have been insufficient to inactivate the agent that causes BSE.

Declan Butler



Supercomputer 'dumping' dispute heats up

Washington. A vote last week in the US House of Representatives has raised the political stakes of a pending decision by the National Science Foundation (NSF) whether to allow the leasing of Japanese-made supercomputers to the National Center for Atmospheric Research (NCAR).

The University Corporation for Atmospheric Research (UCAR), which runs NCAR for the NSF, expects to conclude negotiations "in the near future" to lease four supercomputers built by Japan's NEC Corporation. The supercomputers would be used to run complex climate models. NSF would have to approve any leasing contract.

But the NSF appropriations bill for 1997 passed in the House last week includes a measure sponsored by two House Democrats — David Obey (Wisconsin) and Martin Sabo (Minnesota) — denying salaries to any NSF official who approves the transaction, if the US Commerce Department ultimately finds that NEC is guilty of price "dumping" (see *Nature* 381, 723; 1996).

That verdict will be slow in coming. Last week, Cray Research, which lost the leasing bid to NEC, had not yet filed a formal complaint of dumping to the Commerce Department. But the company does intend to file a complaint, according to William Bartolone, vice-president for government operations, after which the department would have 20 days to decide whether to launch an investigation. If it does, even a preliminary verdict could take four months.

The department could decide to investigate the matter on its own, although that rarely happens. In a letter last week, Neal Lane, the NSF director, asked Mickey Kantor, the Commerce Secretary, to notify him if the department plans such a "self-initiated" investigation or if it finds, after reviewing data recently supplied by the NSF, "that its concerns are allayed so that such an investigation is not warranted".

The NSF has twice asked UCAR to look closely at the NEC bid and to confirm that it

does not include noncompetitive practices. In a statement last week, UCAR said that, after hiring outside consultants to review the bid, it has concluded that "the best and final offer of [NEC] is fair".

The NSF clearly believes, too, that right is on its side. Lined up against it, however, are protectionist forces in Congress. The American Federation of Labor and the presidential gadfly Ross Perot were only two of those who weighed in last week to encourage the House to squash the NEC supercomputer deal.

A recent congressional vote on a Japanese shipbuilding contract, which went in

favour of protectionism, was said to be a factor in a decision by Jim Kolbe (Republican, Arizona) to withdraw his amendment deleting the Obey-Sabo clause.

Kolbe did, however, secure a promise that the matter will be debated again before the House and Senate finalize the NSF appropriations bill this autumn.

That gives the NSF at least several weeks to decide whether to approve the NCAR deal, and to ponder the wisdom of bucking what has been an unwritten rule among large federal laboratories — that, when it comes to supercomputers, the only way to buy is American.

Tony Reichhardt

French action aims to quell BSE fears

Paris. France moved swiftly last week to respond to the advice of its scientific advisory committee on bovine spongiform encephalopathy (BSE), spurred by criticism of its failure in early May to respond to the committee's suggestion to act as if BSE can be passed to humans.

Alain Juppé, the prime minister, announced that the government would implement to the letter a series of recommendations submitted by the committee just a few hours earlier. These include the following:

- Meat and bone meal should be manufactured only from abattoir wastes considered fit for human consumption, and the practice of processing carcasses of farm and domestic animals should be discontinued.
- There should be a ban on the use of all central nervous system tissues of ruminants both in meat and bone meal, and for human consumption. (In April the government banned human consumption of such tissues from animals born before August 1991.)
- France should ask for its ban on the import of meat and bone meal from the United Kingdom to be extended to all members of the European Union (EU).
- There should be strict separation of feed used for ruminants and that used for other animals, enforced by new tracking systems and better controls.

The high political profile given to the announcement appears to have been part of an attempt by the government to restore public faith in its commitment to openness about BSE. This was damaged last month when the newspaper *Le Monde* revealed that the BSE advisory committee had advised the government on 9 May to act as if BSE could be trans-

Alain Juppé (right) announces the French measures on BSE, with agriculture minister Philippe Vasseur.

mitted to humans, but that the government had decided to keep the report secret until 7 June.

The revelations were particularly embarrassing in that they established that Jacques Chirac, the French president, was aware of this recommendation when in mid-May he promised his support to John Major, the British prime minister, for a partial lifting of the EU ban on UK beef products.

The French government's concern to avoid charges of withholding information from the public is also said to have led it to require a group of researchers to hold a press conference last month about a paper, unpublished at the time (and now published in *Nature* 381, 743–744; 1996) reporting that macaques inoculated with infected cattle brain developed similar pathological symptoms to humans with the new variant of Creutzfeldt-Jakob disease.

● The French consumer association UFC-Que Choisir announced last week that it will bring legal proceedings against 'X' for the marketing of potentially contaminated meat and bone meal imported from the United Kingdom (see *Nature* 381, 544; 1996). Jacques Toubon, the justice minister, said that he was considering taking similar action. **Declan Butler**

Former head of cancer charity arrested

Paris. Jacques Crozemarie, who was ousted as chairman of France's biggest medical charity, L'Association pour le recherche contre le cancer (ARC), after a financial scandal earlier this year, was last week arrested and indicted on fraud charges.

Denis Baumont and Michel Simon, the heads of two companies that are alleged to have enjoyed privileged contracts from ARC were also arrested in police swoops and remanded in custody; the two companies are alleged to have overcharged the charity by substantial amounts and taken excessive commissions (see *Nature* 379, 103; 1996). □

Space nuclear power programme at risk

Washington. A joint US-Russian programme to develop a space-based nuclear reactor should either be enlarged to conduct fuelled tests at higher power levels or shut down, according to a report released last week by the US National Research Council (NRC).

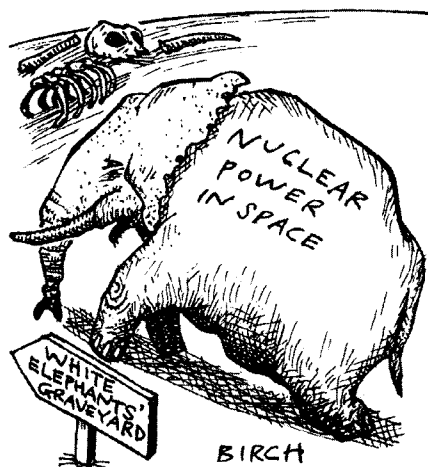
The 30-year quest to develop space nuclear power is in danger of ending, says the panel, unless any of the three US federal agencies involved — the Department of Energy, National Aeronautics and Space Administration (NASA) or the Department of Defense — comes forward with a mission to justify continued research. And that, say most observers, is extremely unlikely.

The United States first flew a nuclear reactor in space in 1965. But since then the research has suffered from uneven support. The 'Star Wars' anti-missile programme of the mid-1980s briefly revived interest, as did NASA's flirtation in the late 1980s with planning human missions to Mars. But by the early 1990s both programmes had been abandoned.

In 1991, the Defense Department acquired the first of several unfuelled Russian TOPAZ reactors, and began working with the Russians on a joint testing programme, which now costs \$8.5 million a year. But the TOPAZ international programme has been hamstrung by political constraints — the

most important being that the project is prohibited from using live nuclear fuel. Instead, the reactor is heated electrically, limiting it to low power levels.

The NRC panel finds no fault with the work of the scientists on the TOPAZ programme, which is run by the Defense Nuclear Agency (recently renamed the Defense Special Weapons Agency). It simply says that, as a result of political constraints, the research is unlikely to produce good information on how nuclear power reactors would work in space.



The panel was chaired by a retired US Air Force general, William Hoover, who once headed the Energy Department's nuclear weapons programme. The panel proposed an alternative to simply scrapping the research. By combining TOPAZ with advanced thermionic reactor research being sponsored by the Defense Department at \$10 million a year, allowing fuelled testing and lifting other constraints, the programme could provide real benefits.

But an even larger stumbling-block is that no agency has come forward with a solid mission for high-power nuclear reactors in space. Uses proposed in the past have included power sources for space-based radars and lunar outposts.

Although NASA studies have concluded that human missions to Mars will need nuclear power, the agency is careful to avoid proposing any grandiose, budget-busting plans. Even the suggestion of using nuclear power in space has become politically charged. "[NASA] seems to be very skittish about anything nuclear," says Hoover. "There's a lot of ambivalence."

So, without support from other agencies, the programme is likely to die. If it does, says the NRC committee, an entire body of knowledge is likely to die with it.

Tony Reichhardt

Budget hits 'political' research funded by tobacco tax

San Francisco. Research considered to have a 'political' aim will not be eligible for support from state funds allocated to tobacco research under a rule adopted during California's budget negotiations last week.

The new budget allocation, which affects \$60.4 million in funding, also overrides the peer-review process and gives a state agency, under advice from a politically appointed committee, the final authority over grants.

The language has aroused the suspicions of tobacco-control advocates and scientists, who say that political pressure may harm funding of research projects critical of the tobacco industry and its supporters in the state legislature.

They say the budget language targets work by Stanton A. Glantz, professor of medicine at the University of California at San Francisco, who has drawn attention to legislators' ties to tobacco interests and to tobacco industry documents which reveal that industry lawyers have kept scientific knowledge of tobacco's deleterious effects hidden (see *Nature* 376, 205; 1995).

It would not be the first time Glantz's work has come under attack. Last autumn, the House of Representatives Appropriations Committee in the US Congress singled out Glantz's work as inappropriate for funding by the National Cancer Institute.

The allocation of the tobacco-research budget follows several years of bickering, in which state agencies fought over how to spend a tobacco tax of 25 cent per pack set aside by Californian voters in 1988 for tobacco-related education, research and health services.

Although the new language restores a large portion of the original funding that voters had directed toward research, a provision would prohibit use of the money for studies "of a partisan political nature", and submit the programme's expenditure plan to the state Department of Finance for review.

The department, in turn, would take advice from the Tobacco Education and Research Oversight Committee, a politically appointed panel which until now has only been generally responsible for the direction of the programme.

Jann Taber, a spokeswoman for the Assembly Speaker, Curt Pringle (Republican, Garden Grove), who inserted the language, said that Pringle's goal was to reinforce constitutional prohibitions against state funding of partisan political research. She claimed that the process of oversight would not change, with the oversight committee acting only in an advisory role. At the same time, however, Pringle's office condemned Glantz's work as inappropriate.

Glantz is reviewing tobacco industry activities in six states. His studies have indicated that state legislators who receive tobacco industry contributions are 42 times more likely to support the industry with their votes than others. According to Glantz's research, for example, Pringle has accepted \$30,750 of tobacco industry money during his legislative career. Under the new rule, the Department of Finance and the oversight committee would decide whether such research violated the rules on partisan political activities.

Mary Adam, of the California and Los Angeles affiliates of the American Heart Association, says that the policy sets a dangerous precedent by allowing a political body to tinker with research.

Larry Gruder, executive director of special research programmes in the office of the president of the University of California, says that the university, which manages the tobacco tax research funds for the state, would object to any attempt to target certain investigators or certain types of research.

Bowing to an outside committee or eliminating work such as Glantz's would violate the university's values and policies, he said. The 1996-97 California state budget is expected to be finalized and approved this week.

Sally Lehman

'Cover-up culture' blamed as officials miss budget hearing

Washington. Two US congressional hearings planned to discuss inconsistencies in the Clinton administration's science budget were postponed last week after the sudden withdrawal of the administration officials who were due to testify.

The hearings would have allowed Republican congressmen to question administration officials about the apparent contradiction between budget projections from the White House Office of Management and Budget (OMB), which imply steep cuts in science funding, and the administration's public pledges to maintain its investment in science.

Robert Walker (Republican, Pennsylvania), chair of the Science Committee in the House of Representatives, which organized the hearings, said that the withdrawal was "one more sign of the culture of cover-up" at the White House. He said he would subpoena witnesses if necessary so that the hearings could take place.

Walker said that officials from OMB and Neal Lane, the director of the National Science Foundation (NSF), who were due to testify on 26 June, as well as Dan Goldin, the administrator of the National Aeronautics and Space Administration (NASA), who was to testify at a second hearing on 17 July, had contacted him "within a few hours" of each other to say that they could not attend. Lane and Goldin said they had prior engagements. But the OMB declined to testify at all, said Walker.

The Department of Energy was also believed to be reluctant to send anyone to the hearings, due to focus on assessments made by the American Association for the Advancement of Science (AAAS) of the effect of the Clinton budget projections on science spending (see *Nature* 380, 572; 1996). The first witness was to be Al Teich, head of science policy at the AAAS.

The AAAS has noted that the administration is projecting steep cuts in NASA, NSF and Department of Energy research spending in 1998, 1999 and 2000. But officials of each agency have said that they are not actually planning such deep cuts. The administration is deeply divided on the issue, and officials at the Department of Energy and NASA say privately that OMB projected the cuts without consulting them.

"We think it is very serious when you have administration officials speaking out of both sides of their mouths on this issue," says Walker. "They should be able to say what they think of the AAAS assessment." He says that the hearings would take place this month or, if necessary, in August, and that the committee may subpoena OMB officials to ensure their participation.

Colin Macilwain

NSF braves threat to abolish social sciences directorate

Washington. The US National Science Foundation (NSF) has established a committee to find a new director for its social sciences directorate, despite the wishes of some powerful congressmen that the directorate should be abolished.

The NSF is seeking a successor to Cora Marrett, who returns this month to her sociology chair at the University of Wisconsin, Madison. But the price of the NSF's defiance became clear last week when the House of Representatives passed an amendment to an appropriations bill that would slash the agency's administrative budget by almost one-tenth.

The amendment was proposed by Robert Walker (Republican, Pennsylvania), chairman of the Science Committee, and passed despite the opposition of Jerry Lewis (Republican, California), who chairs the appropriations subcommittee responsible for NSF funding. The amendment would strip \$9 million from the NSF's \$120-million administrative budget.

Agency officials say it could result in compulsory lay-offs of up to 120 of the NSF's 1,200 headquarters staff in October, when the appropriations bill takes effect. This could pose serious difficulties for the agency, which funds most non-biomedical

university research in the United States. Supporters of the agency say it is one of the most efficient in the federal government, spending only 4 per cent of its budget on administration. But Science Committee members, including Walker and Vernon Ehlers (Republican, Michigan), who spoke in favour of the amendment, are said to be angry with the agency for several reasons.

One is its failure to act on the committee's recommendation that the number of NSF directorates should be cut from seven to six. A provision to do this has been approved twice by the House of Representatives, but has not passed into law. Social sciences, which is the youngest and smallest of the seven directorates, would be the probable casualty of such an adjustment.

But Neal Lane, the director of the NSF, has vigorously defended the social sciences directorate. According to Marrett, the directorate has established social science disciplines as "central" to the NSF since its foundation in 1989.

The Walker amendment transfers the \$9 million from administration to the NSF's research account. But another successful amendment to the bill took \$13 million away from that account to pay benefits to military veterans.

Colin Macilwain

Italian bill sets out recruitment reforms

Rome. Italian universities would be able to select new faculty members from a government-approved list — rather than having to accept candidates allocated to them by the ministry of education and research — under a radical set of new rules being proposed by the new minister, Luigi Berlinguer.

In a bill presented to the cabinet last week, Berlinguer argues that the current system of academic appointment should be abandoned. Winners of the *concorsi*, large national competitions that take place every few years, are at present allocated to a university — usually not of their choice — by the ministry in Rome. The system is widely criticized for its lack of both efficiency and transparency, the latter being seen as allowing appointments to be made on the basis of personal connections rather than merit (see *Nature* 378, 228; 1995).

The draft law proposes that the *concorsi* should continue but that they should take the form of 'procedure di abilitazione'. This means that, like the German *Habilitation* system, they should be used only to establish whether candidates are sufficiently skilled in

teaching and research to be considered for university professorships and associate professorships. Judging committees should include foreigners — a new concept in Italy — while scientific achievement would, for the first time, be assessed "using indicators recognized by international scientific organizations".

Universities would hold separate local competitions for posts as they became vacant, and these would be open to those who have successfully passed the national competition. Candidates would be assessed by committees comprising local and external academics. But the competitions would be closed to those already employed at the university concerned. This, says Berlinguer, would introduce more movement into the university system, and also reduce the influence of personal connections.

Dario Braga, an associate professor of chemistry at the University of Bologna, and a member of the Association for Research and Academic Culture, which has pressed for university reform for several years, says that Berlinguer's proposals are "a big step forward towards honest recruitment".

Alison Abbott

Fears for Portuguese science as Gulbenkian shifts funding focus

London. The Gulbenkian Foundation, Portugal's largest private source of research funding, has decided to reduce its investment in basic research and concentrate on science education and the public understanding of science.

The decision has been criticized by scientists as both discriminating against older researchers — most of those facing redundancy as a result are more than 45 years old — and a blow for research in a country that spends less than one per cent of its gross domestic product on research and development.

But officials at the foundation describe the changes as a "necessary restructuring". They promise to leave open the laboratory facilities at the Gulbenkian Institute of Science, a biomedical research centre at Oeiras, north of Lisbon.

The US\$2-billion foundation is part of the legacy of Calouste Gulbenkian, an Armenian-born British citizen who made a fortune arranging deals between British oil companies and Middle Eastern sheikhdoms between the two world wars. Although most of the foundation's annual spending goes on arts and education, US\$7 million a year currently goes to science.

Under the new arrangements, due to be implemented early in 1998, the 24 researchers at the Gulbenkian Institute will be reduced to a core group of six, concentrating on research in molecular biology, developmental biology and neurobiology. They will also be required to teach and to take part in new activities including communicating science to the public and research methodology to doctors.

"We are not downgrading science," insists João Caração, director of science at the Gulbenkian Foundation. "But we have to be realistic. We cannot be the ministry of science". Caração acknowledges that the advancing years of a number of the institute's staff contributed to the decision to reorganize. But he says any further delay might have posed further difficulties.

Critics such as João Monjardino, a Portuguese researcher who is now reader in cell studies at Imperial College in London, argue that the reorganization will terminate "high quality research", as well as the careers of some high quality researchers. "Portugal has a fragile scientific organization and can ill afford the loss of these scientists," he says.

Caração denies that the foundation is about to abandon the institute's staff. The oldest researchers asked to leave will be offered early retirement, he says. The remainder will be helped to find new jobs in Portugal's state universities. Ehsan Masood

Call for Europe to 'revitalize efforts' in basic research

Munich. Europe needs to reinvigorate its commitment to science if it is effectively to compete internationally, according to the European Science Foundation (ESF), the Strasbourg-based association of national funding agencies from 21 countries.

A position paper on the European Commission's Fifth Framework Programme, due to be launched in 1998, avoids detailed discussion of how the programme might be structured to meet the complex political demands laid out in the Maastricht treaty, or how much money should be allocated to it.

It focuses instead on a scientific agenda that considers both short-term and long-term aims, based on five main themes that it says are crucial to social and economic development in Europe after the turn of the century: information and communications technologies, industrial technologies for complex systems (including, for example, energy systems), molecular mechanisms in health and life, sustainable use of the environment, and a social sciences programme.

In common with the German position paper (see *Nature* 381, 634; 1996), that from the ESF suggests that the commission should delegate the management of some of its research programmes to national agencies. It also suggests a two-stage application procedure, with applicants being required initially to present merely a summary of their proposals, thus saving applicants' time and reducing the burden on reviewing panels.

The ESF also attempts to define its own future role in European research. The foundation, which has been quietly but successfully running scientific networks, workshops and small science programmes for the past 21 years, has recently been trying to carve out a wider role for itself in Europe (see *Nature* 366, 193; 1993), based on the expertise within its scientific standing committees and its wide networks of contacts with basic researchers throughout Europe.

Analyses prepared by ESF standing committees identify many scientific and technological areas that, the foundation says, would benefit from research carried out at the European level. But its position paper stresses that the overall programme can succeed only if a small number of goals are set.

The foundation says that task forces set

up by Edith Cresson, the research commissioner, to promote coordination of themes such as vaccine development across different research programmes, should improve the efficiency with which these goals are met. But it suggests that no new task forces should be created until experience has shown how they can work most effectively.

The position paper says that defending Europe's basic research base will be helped by the continuation of the Fourth Framework Programme's Training and Mobility of Researchers (TMR) programme, which supports fellowships for scientists to work in other countries.

But the ESF says the TMR programme should be expanded to promote the development of specific skills, such as mathematical modelling and algorithms for industrial application. It should also include a return fellowship scheme, similar to that offered by the German Humboldt Foundation, so that scientists can return later to the host laboratory to reinforce the skills learnt.

Peter Fricker, the ESF's secretary general, has asked Cresson for more discussions on how the foundation could help in the development of the Fifth Framework Programme. "We believe that we have a lot to offer", he says.

● Sweden wants the European Union to take special measures to support applications from small and medium-sized enterprises (SMEs) for research grants under the Fifth Framework Programme (FP5), arguing that such organizations lose out because the current system does not adequately consider their needs.

In its position paper on FP5, Sweden points out that SMEs have a vital role in the future growth of the European economy. But, it says, if they want financial support for research and development, they have to spend an excessive amount of time preparing detailed proposals.

Furthermore, the delay between application and start of project is too long for SMEs, given the competitive environment in which they operate. A special SME unit should be created within FP5, says Sweden, to help ease these problems. Such a unit should also be able to transfer applications from one programme to another.

Sweden also suggests that the commission should give greater encouragement to projects with only two or three partners — rather than award grants on the apparent basis of "the more the merrier" — arguing that the current average of five or six partners for each application places an undue burden on project coordinators.

Alison Abbott

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Fricker: offering ESF's services to Brussels.

foto diary

Test ban 'within reach', say negotiators

Paris. After five weeks of tense negotiation, ending last Friday, a treaty banning all nuclear tests seems likely to be sent for approval to the United Nations General Assembly in September. But backroom negotiations in individual countries over the next month will be critical in determining whether the treaty obtains the backing needed to allow its full implementation.

The Conference on Disarmament, which is negotiating the Comprehensive Test Ban Treaty (CTBT) in Geneva, had fixed a deadline of 28 June for drafting a final text, the latest date possible if the text is to go before the United Nations in September.

The failure of the conference to agree on the text last week has been interpreted by some as a failure of the negotiations themselves. But others point out that the conference succeeded in its aim of completing negotiations on the main thorny issues.

Indeed, the text submitted by Jaap Ramaker, the Dutch diplomat who is chairing the conference, on the afternoon of the deadline is expected to be almost identical to the text that will go to the UN, says Rebecca Johnston, of *Disarmament Intelligence Review*.

Only a few weeks ago, the text contained thousands of sections that had still to be agreed. But many contentious issues have since been settled. China has relinquished its demand to be allowed to carry out peaceful nuclear explosions, and all five nuclear weapons states — known as the P5 — have also agreed a ban on low-yield nuclear explosions.

Many delegates had hoped to reach an agreement in principle on the text before the deadline expired. But no formal agreement could have been reached before the text had been reviewed by the government of each country. A further delaying factor is that the outcome of Russia's presidential election will not be known until this week.

The next few weeks will see further intense discussions between government representatives before the last round of negotiations opens on 29 July. Observers are optimistic that no countries will demand major changes to the text, which would end all prospects of having a treaty signed in September. The leaders of the G7 countries reaffirmed their commitment to signing the treaty by this date at a meeting in Lyons, France, last weekend.

The major sticking-points in the negotiations are the conditions under which the treaty would enter into force. The United Kingdom, Pakistan and Russia have insisted that the treaty should take effect only if it is ratified by at least the five nuclear weapons states and the three threshold weapons states, India, Pakistan and Israel — the so-called 'group of eight'.

This has become a Catch-22 situation.

India says it will not sign because the treaty does not provide sufficient commitment to nuclear disarmament, although it says that it will not block agreement by other states. But by making the treaty's entry into force dependent on India's signature, the conference risks scuttling the entire exercise.

Some observers are cautiously optimistic that the discussions over the next four weeks could result in a softening of the conditions on entry into force. The United States said last week that it was willing to be "flexible" on the issue. This has raised hopes that the US administration might exert pressure on the United Kingdom, Russia

and Pakistan to relax their demands.

Eliminating the requirement that the treaty's entry into force should depend on a signature by the group of eight would remove the biggest obstacle to reaching agreement.

Meanwhile, the final text proposed last week also includes a possible way out of the dilemma. Although it insists that the treaty be ratified by at least 45 countries — including the group of eight — before it could enter into force, it says that if this condition is not met within three years, a conference would be convened to find a solution.

Declan Butler

Japan plans for more technicians

Tokyo. In addition to promising major increases in funding, Japan's new five-year plan for science and technology, announced last week by the Council of Science and Technology, could revolutionize employment practices in government research organizations, according to scientists leading the reform of Japan's public sector research system.

The plan was drawn up by the council, Japan's highest policy-making body for science and technology, as required by a new law passed last November. In addition to setting an ambitious target for total government expenditure on science and technology of ¥17,000 billion (US\$160 billion) over the next five years, it calls for the ratio of technical support staff to researchers to be increased to 1:2 for national universities and 1:1 for national research institutes, far above current levels.

Since the late 1970s, Japan's universities have suffered from a steadily growing shortage of technicians, due the government's policy of cutting down on the number of government employees — including university staff. The new policy aims to reverse that trend.

According to a spokesman for the Science and Technology Agency, which helped to draft the plan, this will probably be achieved by hiring part-time or contract employees, circumventing rules that traditionally give lifetime employment to all government employees, including university researchers and technicians. Scientists say that such a move could set a precedent for more flexible employment practices throughout the public-sector research system.

Michio Oishi, director of the Ministry of International Trade and Industry's National Institute for Bioscience and Human Technology, says that this is "the most important plan in 20–30 years".

He adds that among the most revolutionary aspects are the plan's support for contract employment for government and university researchers.

Some institutions — including Oishi's own — are considering introducing systems under which researchers will be granted permanent employment only after their performance has been evaluated over a provisional period of several years. But there is at present no legal basis for the provisional hiring of government employees.

Akito Arima, president of the Institute for Physical and Chemical Research, and chairman of a subcommittee within the University Council that is considering the issue of limited-term employment for university researchers, says that the government is now considering legislation to put such hiring practices on a more secure legal footing.

But although there is general support for more flexible employment practices among government scientists, widespread adoption of such systems will require effective evaluation of individual researchers' performance, says Oishi.

Many observers point out that the Japanese government frequently produces science policy statements that are full of good intentions but are generally not implemented in full. This time, however, despite opposition from the Ministry of Finance, there appears to be a set budget for science and technology which will help smooth the introduction of the promised reforms.

But money may yet prove to be the biggest obstacle to implementation of the plan. With the government facing a huge deficit, the finance ministry's view that the forecast expenditure cannot be afforded may yet still prevail, if political support for the plans should falter.

Stephen Barker

German research council warns of poor career prospects

Munich. Wolfgang Frühwald, president of the *Deutsche Forschungsgemeinschaft* (DFG), Germany's research council, told the annual meeting of the council in Leipzig last week that limited funding has forced the DFG to reduce by almost a half the number of fellowships it can offer in future.

Frühwald said that the move would be particularly hard on young scientists, as the number of research positions in industry has also fallen considerably over the past four years. Germany was in danger of losing a generation of researchers, he said, if students turn away from natural sciences because of poor job prospects.

At the meeting, the DFG approved the setting-up, in cooperation with the National Natural Science Foundation of China, of a research coordination office in Beijing, despite Germany's diplomatic differences with China over human rights in Tibet. □

Scientists 'support Yeltsin'

St. Petersburg. The Scientists' Union of St. Petersburg has published a letter in the local press stating that its members will vote almost unanimously for President Boris Yeltsin in the second round of the presidential elections this week. "If Yeltsin is elected, we will be able not only to strengthen the [academic] freedom we have already obtained, but also increase the influence of scientists' over the destiny of the Russian science," they wrote.

But some Russian politicians take a more cynical view. Galina Starovoytova, a member of the Duma, the lower chamber of the Russian Parliament, has suggested that many Russian intellectuals do not intend to vote at all, as they do not consider any of the candidates an ideal president, and feel that their votes will make

little impact. "As a result the Russian president will emerge without their will or efforts," she said. □

Princeton helps Korean tokamak

Washington. The Princeton Plasma Physics Laboratory in New Jersey is to support a US effort to assist South Korea in the design of a major tokamak facility, which, the Koreans say, will be operational by 2002 at the town of Taejeon. Korea will pay \$540,000 this year for the US assistance, which will also involve researchers at other magnetic fusion facilities in the United States, with additional awards expected to follow.

The tokamak will cost between \$200 and \$300 million, and will use similar technology to the \$750-million Tokamak Physics Experiment (TPX), which was planned for Princeton but cancelled by the US Congress last year. □

Biotechnology grants available

London. The British government is inviting bids for a number of grants of £50,000 (US\$75,000) available to manufacturing and service industry companies that employ up to 500 employees and can devise projects that involve the "innovative use of biotechnology".

The initiative, part of the Department of Trade and Industry's Biotechnology Means Business scheme, is sponsored by *Chemistry and Industry* magazine, with backing from the BioIndustry Association. The closing date for applications is 31 October 1996. Winners will be announced early next year. □

Newsletter for BSE surveillance

Paris. Addressing an issue highlighted by the bovine spongiform encephalopathy (BSE) crisis, a group of European researchers will this month launch a publication aimed at improving European

Even Carl von Linné
would have difficulty
classifying us



collaboration in the surveillance of communicable diseases. According to its first issue, the monthly newsletter, *Eurosurveillance*, will publish data from surveillance networks and the results of investigations into outbreaks. It will also "compare national approaches to communicable disease prevention".

● *Eurosurveillance*, Hôpital National de Saint-Maurice, 14 rue du Val d'Osne, 94410 Saint Maurice, France. □

'Fastest patent in the West'

London. The UK Patent Office claims to have processed a patent application in the shortest time anywhere in the industrialized world. Nick Sills, the inventor of the 'hydrodigger', an underwater dredging and trench-cutting machine that uses water pressure to blast material from the sea-bed, was awarded a patent just 10 months after submitting his application.

The patent was awarded under the Patent Office's new fast-track service, which aims to award patents within one year of the application being filed. The service is intended for small businesses and individual inventors who need a speedy response to help them to secure funding for their ideas. □

Physics prize honours Bethe

Washington. An annual prize for "outstanding and numerous accomplishments" in astrophysics and nuclear physics has been announced in honour of Hans Bethe, the last living giant of the Manhattan Project, and professor of physics at Cornell University, New York.

The \$7,500 prize, which will be awarded each year by the American Physical Society, is supported by endowments from Cornell, Los Alamos National Laboratory in New Mexico, and others. The prize was due to be announced on Tuesday (2 July) at Bethe's ninetieth birthday party at Cornell. □

Sports science prizes

Lausanne. The first US\$250,000 Olympic Prize for Sports Science, sponsored by the International Olympic Committee (IOC) and Parke-Davis, the pharmaceutical company, and open to researchers in the biomedical sciences, psychology and physical sciences, has been awarded to two researchers for their work in demonstrating the relationship between physical activity and the prevention of coronary heart disease.

Jeremy N. Morris, emeritus professor at the University of London, and Ralph S. Paffenbarger Jr, emeritus professor of health research policy at the University of California, Stanford, were selected from more than 30 nominations. They will be presented with the award on 14 July at the opening ceremony of the 105th session of the IOC in Atlanta, Georgia, less than one week before the start of the centennial games. □

New head for Argonne

Washington. Don Eastman, an IBM executive who has a background in condensed matter physics, has been appointed director of the Argonne National Laboratory near Chicago, Illinois. Eastman (left), who is 56, has spent 33 years with the computer company and most recently headed an effort to restructure its hardware development operations. He succeeded Alan Schiesheim, who served as director for 12 years, on 1 July. The multipurpose Department of Energy laboratory, which employs 4,500 people, recently saw its long-term future assured with the opening of the Advanced Photon Source, a powerful X-ray facility (see *Nature* 381, 101; 1996). □



Carl von Linné: 18th century botanist, researcher, physician, professor, lecturer and a resident of the Swedish university city of Uppsala (pronounced OOP-SA-LA). A consummate classifier, Linné systematized the plant, animal and mineral kingdoms as well as drew up a treatise on the diseases known in his day.

If Linné were alive today, he'd be proud of the vast number of diverse and important scientific fields researchers are involved in. Our job is to help life scientists find solutions by getting involved in their activities. We're Pharmacia Biotech—also from Uppsala.

The diversity of our involvement makes us a little tricky to classify at times. After all, our specialists work with such disciplines as (get

ready): cell separation, cell culture, nucleic acid purification, PCR-related areas, cDNA synthesis and cloning, vectors, nucleic acid electrophoresis, hybridization, sequencing, gene expression, restriction enzymes, modifying enzymes, nucleotides, oligonucleotides, polynucleotides, oligonucleotide synthesis, chromatography media, reagents, BioProcess chromatography, electrophoresis systems, electrophoresis media and spectrophotometry. As you can see, our company works with more than just one scientific discipline.

Do you?

If so, give us a call: 1 (800) 526-3593 in the USA; +81 3 3492 6949 in Japan; or +46 18 16 50 00 in Europe and the rest of the world. Or visit us on the Internet: <http://www.biotech.pharmacia.se>.



Uppsala, Sweden. (And the rest of the world)

Swifter, higher, stronger: pushing the envelope of performance

Although athletic records continue to be broken, constraints on human physiology inevitably imply overall limits to achievement. Approaching these limits may change the nature of the Olympic Games.

In two weeks' time, the Olympic torch will arrive after its 3-month journey from Greece to light up the centennial games in Atlanta, Georgia. But as the flame opens the games, the motto of Baron Pierre de Coubertin, the French founder and inspiration behind the modern Olympic games, that "the important thing in the Olympic Games is not winning but taking part", is likely to find few takers among modern contenders.

Today's athletes prefer the modern Olympic motto, *citius, altius, fortius*, adopted in 1920, which translates as 'swifter, higher, stronger'. Olympic performances have been improving steadily since the first of the modern games in 1896 (see figure, right). No world record achieved before 1980 survives to Atlanta. Even the US athlete Bob Beamon's long-jump record of 8.9 metres — thought to be unassailable — was broken in 1991 after 23 years. And at the end of last month, Michael Johnson clocked the fastest 200 metre sprint, breaking a record that had held for the past 17 years.

But de Coubertin's motto may not be entirely redundant, although for different reasons than he could have foreseen. Many researchers do not expect the trend of continuous record-breaking to last indefinitely. They claim that there are limits to the strength, speed and height that the human body can attain. If so, it suggests that a diminishing ability to break records will, in future, mean that human athletes will have to compete for the sake of competing.

Roger Woledge, for example, professor of physiology at the Institute of Human Performance at University College, London, is one of those who believes there is an "absolute limit" to athletic capability. "There is a direct analogy with racehorses," he contends. "Methods of breeding keep changing,

but there is still no surge in records."

The situation with racehorses, he says, can be explained by the fact that these horses in Britain derive from a small genetic stock, namely just three animals that were introduced by the Arabs at the beginning of the eighteenth century. Humans are far more diverse, he admits. "But the principle of a performance limit still holds."

This thesis, however, is not shared universally by physiologists. Many agree that a limit to human performance may, in the-

One person who takes this line is Jared Diamond, professor of physiology at the medical school of the University of California, Los Angeles. Diamond has compared the metabolic rates as indicators of energy production of 37 species — including humans — both at rest and after exercise (see *Proceedings of the National Academy of Sciences*, 87, 2324-2328; 1990). He found that in most cases, the ratio did not exceed 7:1.

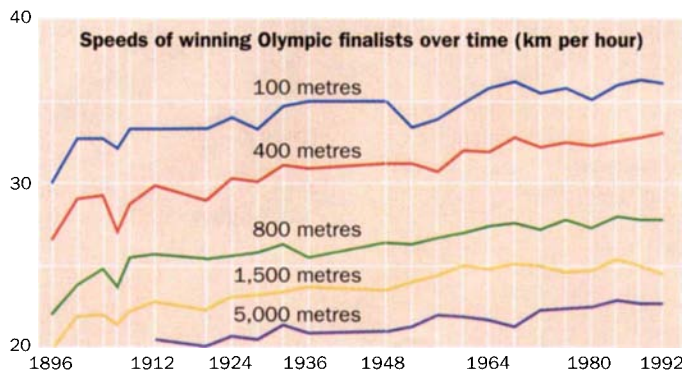
This was a surprise, as the metabolic rate for individuals at rest varied up to 30,000-fold between species. "Human experience shows that even motivated athletes in training, soldiers and explorers with unlimited available food cannot maintain chronic activity levels much beyond 40,000 kJ per day," says Diamond, himself a former amateur weightlifter, basketball player and middle-to-long-distance runner. "I can't see records going up forever."

Peter Keen, senior lecturer in sports science at the University of Brighton, agrees. The coach of world-class cyclists such as the former men's world pursuit champion, Chris Boardman, and the current women's champion, Yvonne McGregor, Keen says that continued increases in world records are not a reliable guide to the future of athletic performance. "There is a limit to the extent you can pump oxygen round the lung, and lungs aren't going to get any bigger," he says. "Most of our best athletes are already operating close to the maximum."

Élite but atypical

Most researchers agree that improving physiology through training — whatever its limits (see page 14) — is just one component of improved athletic performance. Today's athletes are also trained in confidence-building, and are selected from a much wider section of the population than their predecessors. Both factors have helped to push up records on track and field.

But Keen points out that the advantage of selecting elite athletes from a wider gene pool may soon diminish, as average levels of fitness in the developed world have begun to recede. "Elite sportsmen and women are not a fair reflection of society," he says. "We're fatter, more sedentary and less fit than we used to be, and this has implications



Going up: athletes' speeds in the Olympic finals show continuous improvement over the last century. How long will this last?

ory, exist. But they point out that there is still inadequate quantitative data indicating whether this hypothesis can be confirmed.

Others believe that improvements in sports equipment, clothing, nutrition and 'motivation techniques' will compensate for any potential levelling off in human physiological performance. Meanwhile, a third group does not see any reason to believe in performance limits at all.

What all three groups agree on is that adequate data to enable comparison of the rival hypotheses is lacking. "The question of performance limits is moving away from science and into the realm of 'chit-chat'," says Robert McNeill Alexander, professor of zoology at the University of Leeds. "The existence of a theoretical possibility [for limits to human performance] is not a very useful statement. We need hard data and a hypothesis we can test."

Yet despite this lack of data, some physiologists still argue that our existing knowledge of the broad limits of the physiological performance of various human organs suggests that humans — particularly those engaged in endurance activities — are probably approaching a performance ceiling.

On other pages

- Trawling the gene pool 13
- The trials of training 14
- Women catching up on men? 15
- The drug-use dilemma 16

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Carl Lewis (left): the only man to equal Jesse Owens' four Olympic gold medals on the track. for the future of athletic performance."

Some take this further. In *Why Things Bite Back: Technology and the Revenge of Unintended Consequences*, published earlier this year, Edward Tenner, a former physical science and history editor at Princeton University Press, points out that rising living standards and a suburban environment are "probably the reason for the inability of

today's top [baseball] pitchers to match the velocity and skills of the stars of earlier generations".

Supporters of such arguments point out that, at the 1936 Olympic games in Berlin, the US athlete James 'Jesse' Owens, who, along with Carl Lewis, is the only man to win four Olympic gold medals on the track in the same games, clocked 10.3 seconds in the 100-metres sprint, wearing leather shoes and running on a track filled with compressed cinders.

Nevertheless, the contribution of improved — and safer — sporting equipment to world records cannot be underestimated. Impact-absorbing running shoes and smoother, more resilient tracks; aerodynamic javelins and fibreglass poles and rubber foam cushioning for pole vaulters; these have all contributed to improved performance.

Between 1942 and 1960, for example, the world record in the pole vault improved by only 5 cm. But after the introduction of a

fibreglass pole in 1963 the record increased by 23 cm in the first year and has since improved by a further 132 cm.

The unpredictability of limits

Such advances in equipment fuel the arguments of those who claim that the rate of progression in world records is not about to decline. Edward C. Frederick, for example, a former director of the Nike Sport Research Laboratory in the United States, is also among those who believe that the importance of psychology to human performance is often underrated in predicting future achievements.

Frederick is a former US college athlete who ran the mile in four minutes 17 seconds in 1968, and is now a director of Exeter Research, a sports footwear research consultancy in Exeter, New Hampshire.

He says that scientists should stop "embarrassing themselves" by trying to predict the limits of human performance. "Ever since records have been kept, they have been inexorably broken." **Ehsan Masood**

Bannister urges spreading the net

London. Sir Roger Bannister, the first man to run a mile in less than four minutes — a performance described as breaking "the barrier of barriers" — is no stranger to the limits of athletic performance.

Two years before his historic run, he had been widely tipped to win the 1,500 metres race at the Helsinki Olympics of 1952. Coming into the final stretch of the race, he was, he says, "perfectly positioned with a free run ahead of me. But the finish was just not there."

He finished fourth, and the diagnosis was straightforward. "Muscle problem was what led to my failure," says Bannister, who subsequently became one of Britain's leading neurologists, and stepped down in 1993 as master of Pembroke College at Oxford. "I had trained very little — about half an hour a day, with about five races a year — and the muscles had got used to that. But in Helsinki I was suddenly confronted with three races in three days."

Last year, Bannister stirred up controversy in the British press with remarks to the British Association for the Advancement of Science that black athletes may have certain 'natural anatomical' advantages. These, he suggested, might include a better power-to-weight ratio due to the relative lack of subcutaneous fatty insulating tissue related to living in a warm climate, and even a longer heel bone.

Perhaps inevitably — and despite his own caveats — the statement gave rise to charges of genetic determinism. But Bannister is the first to admit that the ability to break records depends on a complex set of factors, some biological, some psychological and some cultural. "In many run-

ning events, physique is not particularly important," he says. "But physiology is — and mental drive is more important than either of these."

In some cases, the physiological differences, even within human races, seem innate. Muscle biopsies from runners have shown that sprinters tend to have a relatively higher percentage of fibres undergoing anaerobic respiration, limiting their need for oxygen over a short burst.

But even innate characteristics are susceptible to external factors, such as training at high altitude. In recent years, runners from Kenya have dominated middle- and long-distance events at the Olympics. The reason may be as much environmental as genetic. "If you live at 8,000 feet, with no school bus service, by the age of seven or eight you are covering large distances every day," he says.

Bannister points out that variation of the human gene pool in terms of athletic performance is fairly widely distributed across all populations. "So far, it has not been proved that members of any one group are so consistently and repeatedly superior as to remove the chance of others beating them," he says. "There was a time when we thought that no-one could beat a Kenyan runner again; but that has not been the case."

One implication, he suggests, is that future improvements in performance will come less from harder training than from selecting athletes from a wider gene pool. "You need an efficient means of selecting those early on whom you think have talent. You then give those with talent expert

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Times gone by: 4-minute record is still a milestone.

coaching — including a global look at how their lives may develop — as well as a back-up physiology service, able to recognize when they are being overtrained."

Last month, Bannister proposed to the British government that it should set up a special sports scholarships fund, based on funds raised through the new National Lottery, to help do precisely that. While acknowledging that the margin by which records are being broken is diminishing, he confidently predicts that a three-and-a-half minute mile will be run within the not-too-distant future.

Science, he suggests, can help to improve performance considerably, not least by helping to ease some of the stress of training and competition; many athletes, he says, fail to reach their full potential because of a lack of the necessary mental skills, including the ability to overcome the pain of extreme effort.

But there will always be factors that science cannot reach. Bannister points out that when Jonathan Edwards — a physicist by training — set a new world record in the triple jump last year, "he did not know how he had done it." □

Science tracks down the training dangers

EXERCISE, as many of today's athletes are discovering, can be bad for you. Moderate exercise has a beneficial effect on health. But at the levels maintained by dedicated athletes it can have various harmful side-effects. Increasing demands on the body have begun to expose the subtle physiological costs of chronic exercise. Understanding the nature of these costs has become central to high athletic performance.

There are several reasons why the negative impact of excessive exercise has become increasingly prominent. Perhaps the most important is the fact that improvements in sporting achievements over recent decades have been the result of ever more demanding training regimes. Club runners 30 years ago might have expected to cover 20–25 miles a week; but their modern counterparts may cover three times that distance.

Observations of overtraining

In the past, knowledge of the dangers associated with high levels of exercise and how they can be overcome was derived chiefly from the practical experience of individual athletes and their trainers. More recently, an increasingly scientific approach to all aspects of sport means that a more systematic approach has been adopted to studying the physiological effects of intensive training.

As a result, it is now becoming apparent that, rather than presenting intractable limits on the levels of performance achievable by the human body, such phenomena are seen more as dangers that trap unwary athletes who fail to take account of physiology when drawing up their training schedules.

The most easily understood — and earliest recognized — conditions resulting from hard and continuous training are 'over-use' injuries due to the failure of muscles, bones or joints under a repeated, heavy load. For most adult athletes, such injuries, which include stress fractures of bones,

strained muscles, snapped tendons and inflammations of joints (such as 'tennis elbow'), are inconvenient, but can be overcome with time.

For children, whose bones are still growing, the consequences can be more serious. Adult success in sports such as swimming and gymnastics can require hard training even before an athlete has reached teenage years. One study of young female gymnasts found that 11 per cent already had fractures in the vertebrae of the lower back, four times the occurrence in the normal population. Fortunately, potentially disastrous injuries can be avoided through the careful planning of training schedules.

The most dramatic training-related syndromes concern the cardiovascular and respiratory systems. Sudden death due to heart failure is uncommon but not unknown. Indeed, the first marathon runner, Pheidippides, expired after his run from Marathon to Athens in 490 BC with the words "rejoice, we won" — although Dan Tunstall-Pedoe, medical director of the London marathon, has seen only four deaths during the race's sixteen-year history.

Such sudden deaths are usually the result of training exacerbating one of a number of congenital heart defects (although heart failure in professional cyclists with no such complicating factors has been known). The most common cause of sudden death in young athletes in the United States, for example, is due to a condition known as hypertrophic cardiomyopathy, in which the walls of the left ventricle of the heart become abnormally enlarged. The natural thickening of the heart muscle due to exercise further reduces the size of this chamber, resulting in a fatal heart attack when the heart is placed under extreme strain, as during competition.

Asthma, the chronic constriction of the bronchial airways, is another potentially fatal condition that is aggravated by training. Although no less serious than congenital heart conditions, its effects are less dramatic. It is also far more common.

Asthma and other effects

According to Mark Harries, Clinical Director of the British Olympic Medical Centre, between 10 and 15 per cent of athletes, even at the highest level, report occasional symptoms of exercise-induced asthma. As with its rise in the general population, the factors that aggravate asthma in athletes are far from clear, for example cold, dry air induces asthma more effectively than warm, moist air. But it can be easily treated by inhalation of bronchodilators, some of which are permitted for use by athletes.

Some of the deleterious effects of training for high-level competition are gender-

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Testing times: an athlete during physiological tests.

specific. In particular, the reproductive physiology of women athletes is very sensitive to disturbance under extreme training regimes. Roger L. Wolman of the Royal National Orthopaedic Hospital in London has estimated that in sports such as cycling, running and rowing, more than half the women competing at international level have disturbed menstrual cycles (oligomenorrhea). Similarly, more than 60 per cent of women gymnasts do not menstruate (amenorrhea), and delayed onset of menstruation is common.

Exercise is only one factor

These dysfunctions may or may not be related to sport, and are not caused exclusively by exercise. Thus although an athlete's training is one of a number of contributing factors, pressures to maintain a specific physique — for example, light weight in rowing coxes or female gymnasts — can produce similar symptoms to those of sufferers from eating disorders such as anorexia.

The cause of athletic amenorrhea is unclear. One theory, of which Hans Kreitzer of the University of Limburg in the Netherlands, is a leading proponent, is that elevated levels of endorphins and cortisol, due to physical and emotional stress, act on the hypothalamus, interfering with its secretion of gonadotrophin-releasing hormone, and thus suppressing levels of gonadotrophins.

Most seriously, the low levels of oestrogen during amenorrhea result in a reduction in bone density which, Wolman and others suggest, can leave the athlete vulnerable to osteoporosis in later life. Reduced gonadotrophin levels have also been recorded in male marathon runners — though much less frequently than in women — and have been associated with low sperm counts.

There is much anecdotal evidence that athletes suffer from an abnormally high incidence of colds and influenza. Only recently, however, has it been conclusively demonstrated that extremes of exercise do indeed reduce the effectiveness of the immune system. In one study, for example, B. K. Peterson of the Copenhagen Muscle Research Centre showed that exercising to exhaustion reduces the level of T lymphocytes and natural killer cells for up to 20 hours in both

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

At the limits: carefully monitored preparation can help avoid failures in performance.

athletes and untrained individuals.

Similarly, Laurel MacKinnon, of the University of Queensland, Australia, has reported that antibody levels in saliva are reduced by 70 per cent in racing cyclists after two hours of hard pedalling. And studies by E. D. Bateman at the University of Cape Town, South Africa, have shown that the fastest runners in marathons and ultramarathons have their lowest white blood cell counts and highest incidence of respiratory infection in the days following races.

As with amenorrhea, these immunodepressive effects seem linked to increased levels of stress hormones. In addition, Richard Budgett, medical director of the British Olympic Association, suggests that some athletes engaged in heavy training are unable to sustain levels of glutamine, which is produced in muscle and is essential for the metabolism of lymphocytes. During exercise, muscle tissue uses up the glutamine, effectively starving the white blood cells.

Avoiding chronic fatigue

The suppression of the immune system due to prolonged endurance training does not seriously compromise long-term health but does have serious consequences for performance. The British athletes Diane Edwards and Seb Coe are only two of the many well-known examples of athletes who have failed to achieve their potential at international events due to infections — in both cases toxoplasmosis, usually found as an opportunistic infection of immunodeficient individuals.

Intimately associated with a reduction in general health is the dramatic onset of chronic fatigue experienced by some athletes during periods of particularly intense training, or when suddenly changing training regimes. This so-called 'over-training syndrome' is thought by some to be related to postviral fatigue syndrome (myalgic encephalomyelitis, or 'yuppie flu') and is characterized by reduced performance, fatigue and depression.

Such a condition is particularly insidious for athletes as its treatment — light exercise only slowly increased up to normal training levels — runs counter to the common misconception that harder training always leads to better results. Its underlying causes are still very poorly understood.

A fall in the density of red cells in the blood may be involved, and can be countered by altitude training (or illegal blood doping). In a study of US college swimmers, about 10 per cent of whom are chronically fatigued, W. P. Morgan, of the Kinesiology Department at the University of Wisconsin-Madison, showed that monitoring of athletes' mood and resting heart-rate helped to predict, and thus help prevent, its onset.

Usually, however, over-training syndrome strikes with little warning and can be overcome only by a period of two or three months of light training, which can ruin an athlete's preparation for a major event — such as the Olympics. **Christopher Surridge**

Could women take a lead over men in the long run?

ARE there sports in which women might one day out-perform men? As traditional views about the physical limitations of women change, the question no longer seems absurd. This is particularly so since, after a comparatively late start, women have made dramatic improvements in performance in almost every competitive sport over the past 50 years.

The fact that this rate of increase has been faster than that for men — whose competitive efforts started earlier — over this period means that, based on a simple forward projection of record achievements, women would overtake men early in the next century (see *Nature* 355, 25; 1992).

In practice, we are seeing the upward slope of a S-shaped curve that will eventually flatten off. "Whether the rates will slow abruptly or slowly, we don't know, but they will slow," says Brian Whipp, professor of physiology at St George's Hospital in London, who has been closely studying the statistics.

But it is already possible to point to some sports — particularly those requiring high levels of endurance — in which women may have the potential to outperform men.

For most sports, particularly those highlighted in the Olympics, such a possibility is made highly unlikely by two major physiological differences between the sexes: the proportion of body fat, and the mass of muscle. Speed on the track requires muscle power, and the male sex hormone testosterone increases muscle mass. Conversely, female sex hormones increase body fat — but fat, to an athlete, is just excess baggage.

But the factors that boost and limit performance in some less mainstream events, particularly those described as 'ultra-endurance', are less well-defined. These include the longer triathlon races, which involve a 3.8-km swim, followed by a 180-km cycle ride and then a 42-km run, as well as various other events (such as long-distance swimming) that are not included in the Olympics.

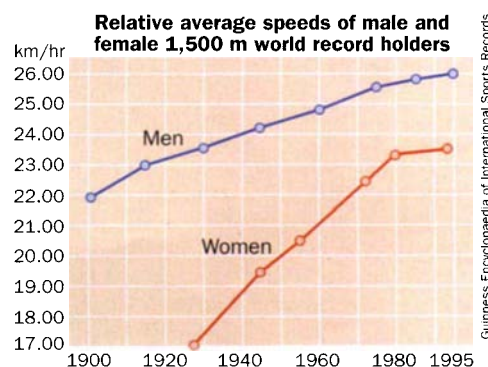
Women's potential in such events has come as something of a surprise. In the past, for example, they were considered to be physically incapable of running in conventional long-distance events, such as the marathon. Indeed, such a belief meant that the first women's Olympic marathon was not held until 1984.

The belief appeared to be given scientific backing by studies in which physiologists claimed to have demonstrated that women were less able to tolerate heat stress. But most such studies had a fundamental flaw, namely that they essentially compared the

physiological characteristics of sedentary women to those of athletic men. The reason was simple: there were relatively few highly trained female athletes available for study.

About 20 years ago, however, in sharp contrast to prevailing ideas, some physiologists and others started to speculate that women might actually be better suited than men to endurance events. The reason, they suggested, was precisely the fact that women have a higher proportion of body fat.

In long-distance swimming, extra body fat is certainly a distinct advantage, providing both extra buoyancy and insulation. Penny Lee Dean, for example, held the world record amongst both men and women for



swimming the English Channel for 16 years, up to 1994. And Lynn Cox successfully swam from Alaska to Russia — a feat previously deemed impossible by physiologists because of the enormous distance and extreme cold.

More controversial was the suggestion of a different type of advantage in running events. During endurance races, an athlete uses both fat and carbohydrate (in the form of glycogen) to fuel their muscles, and the ratio can be critical. By using more fat to provide this fuel, athletes can preserve their glycogen stores and so continue for longer.

Glycogen or lipid metabolism?

It is only when glycogen runs out that an athlete reaches the limits of his or her endurance. Women's extra fat, it was suggested, might mean that they are less reliant on glycogen metabolism, and can make better use of lipid (fat) metabolism.

Subsequent studies ruled out this claimed advantage of women, showing both sexes mobilize fat in much the same way, and the hypothesis has generally gone into disfavour. "The suggestion that women might be at an advantage because they have a higher body fat than men was absurd," says Ben Londerree of the department of health and exercise sciences at the University of Missouri. "It was like saying that fat men should perform better than leaner men." ►

female athletes who had recorded identical times for the marathon (42 km). When the same athletes ran 90 km, the women significantly outperformed the men.

In such events, "factors that we know are important in sprint events don't apply", says Lynn Fitzgerald, an immunologist and reader in sports medicine at Brunel University in west London, and herself a former world-record holder in several ultramarathon events.

Fitzgerald's own preliminary studies, carried out with a research student, Stuart Mackie, suggest that such gender-based differences in performance at very long distances are indeed linked to the respective ratios of fat and glycogen metabolism. "The physiological and psychological traits that are important for

long-distance events are not related to testosterone," she says.

Endurance training is well-known to have a 'glycogen-sparing' effect, raising the consumption of free fatty acids. But individuals with an inherent ability in endurance events may naturally use more fatty acids. Such individuals — who may turn out to be more often women than men — must work harder, as fat metabolism is less efficient than glycogen metabolism. But they are also able to keep going longer, because their glycogen stores take longer to deplete.

This possible gender difference could, according to Fitzgerald, explain the fact that, although only one in four competitors in ultra-endurance races are women, they are often found in the top 10 per cent of those who finish. It remains unlikely that a woman would eventually win a race of this kind; their absolute physiological capacities are smaller than those of men. But such a result is no longer inconceivable. **Ayala Ochert**

Outstanding: the swimmer Alison Streeter ends her 32nd Channel crossing, a record for both men and women.

► But recent studies that concentrate on elite athletes are finding that there may still be gender differences in the way muscles are fuelled in, for example, very long-distance running. David Speechly and his colleagues at the University of Witwatersrand in South Africa, for example, studied male and

Performance-enhancers pose dilemma for rule-makers

MARK HARRIES, of the British Olympic Medical Centre at Northwick Park in London, compares the human body to an internal combustion engine. Both translate chemical energy into mechanical energy; both produce carbon dioxide and water vapour as waste; and both can be made to achieve high levels of performance — through turbocharging in the case of engines, and performance-enhancing substances in the case of athletes.

Drugs have long been used to compensate for the physical and physiological limits of the human body. Their excessive use is widely considered to conflict with the Olympic spirit, artificially distorting any 'natural' difference in ability between competitors. Yet the limits of what is acceptable has never been clearly defined. And the modern Olympics have been plagued not only by efforts of some athletes to disguise their use of illegal substances, but also by debates over the boundary between legitimate and illegitimate ways of boosting performance.

There is an entire chemist's counter of formulations that can achieve this. Stimulants such as amphetamines, ephedrine, cocaine and help to increase alertness, reduce fatigue and generally boost competitiveness and hostility. Narcotic analgesics help athletes to manage pain better. While anabolic agents — such as formulations containing testosterone — can increase muscle strength and bulk, and promote aggression.

One problem of determining what is acceptable is that not all techniques are based on the use of artificial chemical substances. For example, whereas the delivery of oxygen to the muscles is modestly enhanced by training at high

altitudes, where the rarefied air increases the proportion of oxygen-carrying haemoglobin in the blood, larger increases in haemoglobin can be obtained from last-minute transfusions of red blood cells — as demonstrated by the gold-medal winning US Olympic cycling team at the Los Angeles games of 1984. An alternative is to inject erythropoietin, a hormone synthesized by the kidney that stimulates bone marrow to make red blood cells.

In general, any drug — a definition which includes all those substances listed above — is banned by the International Olympic Committee (IOC) if its use either leads to excessive performance enhancement or poses a risk to the health of an athlete. High doses of stimulants, for example, are known to lead to elevated blood pressure and increased and irregular heart beat. Analgesic narcotics carry a high risk of both physical and psychological dependence. Anabolic steroids decrease the size of the testes in males, and lead to loss of breast tissue and diminished menstruation in females. The use of hormones can also cause problems, such as the link between growth hormone and Creutzfeldt-Jakob disease.

In recent years, sports authorities have decided to crack down hard on drug use by lengthening the list of banned items. But this has itself caused new dilemmas, as many items on this list contain substances that are commonly found in over-the-counter medicines, such as remedies for colds, flu, asthma and hay fever, which an athlete may have been taking.

Rodney Walker, chairman of the United Kingdom Sports Council, acknowledges the difficulties involved. But he says that the council's first duty is to eradicate drug-

taking. "Striking the balance between protecting those who don't cheat from those who do, and between the use of permitted medications and the abuse of therapeutic treatments by those not suffering an illness or injury, is not easy," he says. "If this means that we appear to be creating an unfair situation for [sports] competitors, so be it."

Another dilemma is created by the fact that, as fast as sports authorities refine their ability to detect banned substances, researchers are continually coming up with novel ways of raising performance. The latest example is the development of increasingly sophisticated formulations that blur the distinction between performance-enhancing stimulants and food supplements taken in concentrated form.

Two committees of the British Olympic Association (BOA), for example, are currently divided over the use by athletes of creatine, a naturally-occurring substance — it is present in meat and fish, and is also used as a food additive — that is used primarily to help the muscle produce energy, and can thus improve performance in events needing explosive power.

Creatine use is legal. Athletes, such as Colin Jackson, the British holder of the 110 metre hurdles world record, use as much as 20 grammes daily, ten times the natural dietary amount. Paul Greenhaff, a senior lecturer in physiology at the University of Nottingham, and a member of the BOA's nutrition committee, sees no difference between creatine use and carbohydrate loading, which no-one disputes. But Harries, a member of the BOA's medical committee, says it should be banned until the longer-term consequences of such high doses are known. **Ehsan Masood**

Ownership of human genes

SIR — Ownership by capture is an elementary principle in English and American law. Where ownership in a thing does not already exist, legal title is awarded to the person who is first in time to possess it. This rule has been liberally and zealously applied, giving the captor of the undomesticated animal and the extractor of natural waters good title against all the world. It is inevitable that this same canon of the law should find its way to the genes (*Nature* **381**, 11–14; 1996). Who else should own the gene but the scientist who first captures it by cloning?

In an early American lawsuit, Post took Pierson to court in New York for interfering with his right to a wild fox. Pierson had intercepted Post in pursuit of a fox, killed it, and carried off the bounty for himself (*Pierson v. Post*, Supreme Court of New York, 1805, 3 Cai. R. 175, 2 Am. Dec. 264; J. Dukeminier & J. E. Krier, *Property*, Little, Brown, pp 15–19; 1988). The court, relying on English and Roman law, wrote that Pierson as the captor had good title to beast *ferae naturae*, wild and free. Until its status was changed by Pierson's dominion over it by actual possession, the fox was available and unowned. The court, for policy reasons, judged that the fox was a pest to farmers so it was for the good of society that fox hunting be encouraged. Capture as ownership made the outcome of hunting events more certain.

In later suits about property rights of natural resources, the wild and migratory nature of underground water, gases and oil evoked the image of Post's fox. Courts adopted the analogy, and possession, again, became the touchstone of ownership. It made sense to modern jurists considering the increased value of the thing *ferae naturae* once it had been captured and reduced to possession.

Genes, like wild animals and underground waters, are a wandering, flowing resource, acted on by the physical forces of nature. Individuals in a species are linked by their genes. There is a recurrent exchange of genes between populations and the individuals who comprise the populations. The sum of all the genes of the species is the gene pool. Much as the Earth stores water for human consumption, the gene pool preserves the population's genes for future times. It is a genetic repository. If, for example, as a result of a sudden environmental change or the appearance of a new parasite, a formerly neutral allele were to confer a benefit on individuals having it, the frequency in the population of the allele would increase, as it flowed from the gene pool to the individual.

A gene's status in nature is no different from the wild beast or migratory waters. Until cloned, it is gene *ferae naturae*. The

value is in its taking. Gene hunting should be encouraged for policy reasons, because it leads to research and development of products that enhance public health.

Ownership by cloning is reasonable. The alternative would be too uncertain. A claim by one person to one gene as it exists on his chromosome is likely to be contested by another who believes he is the owner because it exists on *his* chromosome. At the very least, siblings will share some of the same genes and thus will be forced to battle among themselves as to who is the rightful owner. Moreover, even if siblings had a common goal and were willing to settle on joint ownership, the controversy would not end there. Shared genes, of course, extend in all directions, familial, geographical and temporal. As they exist on chromosomes, genes cannot be considered personal property. If gene ownership were decided by who in the population had a chromosome having a particular allele in the gene pool, who would get paid? Only those who have it? Or all contributors to the gene pool as conservators of the genetic material? Can you buy the right to a gene in future offspring?

Richard M. Lebovitz

Milen, White, Zelano, & Branigan,
Arlington Courthouse Plaza I,
2200 Clarendon Boulevard, Suite 1400,
Arlington, Virginia 22201, USA
e-mail: Lebovitz@MWZB.com

SIR — The Commentary article by Thomas *et al.*, "Ownership of the human genome" (*Nature* **380**, 387–388; 1996), is valuable, but the interpretation of the findings should be approached with caution.

Patents do not confer "ownership of the human genome" in an expansive sense to an individual or organization; rather, they provide a legally enforceable exclusive right to make, sell or use the patented invention for a limited period. With this in mind, we are undertaking an analysis of US patents, starting from a base of 1980–93 patents selected by the US Patent and Trademark Office for the congressional Office of Technology Assessment (OTA). This database was turned over to the Kennedy Institute of Ethics at Georgetown University in Washington DC so that the institute could catalogue the collection and make the patents available to the public. The institute plans to update, catalogue and begin to analyse the database and expand it to include foreign patents.

This collection of patents was also selected to identify "human DNA sequence patents", but we have found in our preliminary analysis that only a few patents produced by a simple database search strategy correspond to what most molecular biologists would regard as human genes or even

parts of the human genome. Some of the most important and valuable such patents are for genes and gene products — enzymes, peptide hormones, receptors and other obviously relevant macromolecules. But many patents cover production methods, research methods, veterinary vaccines based on traditional biology, sperm and oocyte selection methods, and even peptides that contain the word 'gene' in their name for historical reasons but lack claims to DNA sequences or genes. A large number of the patents captured by a search that combines 'gene' and 'human', 'DNA' and 'human' or 'human' plus a DNA genetic initiator string ('AUG' or 'ATG') capture genes from any organism as well as some peptide and methods patents. It would be useful if authors specified how they defined "patented human DNA sequences".

Moreover, the real work of patents — what limits others' rights — is done in their claims, not what is disclosed about the invention. We are attempting to categorize more precisely the nature of DNA-based inventions and the scope of their patent claims, to enable more refined analysis of patent subcategories. It might be useful, for example, to identify and compare only those patents that claim full-length genes.

Differences in patent strategy can also make patent counts misleading. Many patent applications filed first in the United States, for example, contain numerous broad claims. The Amgen patent for erythropoietin, arguably the most valuable human DNA patent so far, is quite long and has dozens of broad claims for the gene-containing vector, use of the gene, production and use of the gene product, and cells transformed by the vector (the corrections alone run to more than 40 pages). A beta-interferon patent, in contrast — first filed in Japan on behalf of Kyowa Hakko Kogyo — contains only a single claim for the expression vector, with no claims about the gene product, the transformed cell lines, or uses of the gene product or vector. Yet both patents are counted equally if the only index is patents issued. A broad patent of the former type thus might cover an expanse of intellectual property comparable to that of five or six narrower ones, as Thomas *et al.* point out.

This is not a minor problem, but one that can severely warp interpretation of patent counts, as the differences in scope among patents are demonstrably large, and differ systematically. This problem is quite severe when comparing patent applications initiated in the United States to those originating in Japan. Raw patent counts can be analytically useful, but must be treated with caution, and in the case of DNA-based patents, analysis of patents grouped for similarity in the scope of claims and subject matter may be required.

Finally, the financial value of patents ►

► appears extremely variable. Multi-billion-dollar biotechnology patents, such as the Amgen patent for erythropoietin and the Genentech patents on tissue plasminogen activator, are the rare exceptions and not the rule. Even patents of intermediate value, such as those covering broadly useful research and production methods are unusual (for example, the Cohen-Boyer patents for recombinant DNA, the Caltech and Applied Biosystems patents on automated DNA sequencing and pertinent chemistry, or the Cetus patents on polymerase chain reaction). Most patents appear to have vastly less cash income value, and many seem unlikely to tap into large commercial markets at all. When analysing commercial uses of genome research, one cares mainly about the most economically valuable patents, or patents that might constrain research, improvements, and development by others — such as those covering broadly applicable research methods or macromolecules relevant to research with diverse applications.

Robert Mullan Cook-Deegan

National Academy of Sciences,
2101 Constitution Avenue NW,
Washington, DC 20418, USA

LeRoy Walters

Doris Goldstein

Stephen McCormack

Kennedy Institute of Ethics and National
Reference Center for Bioethics Literature,
Georgetown University,
Washington, DC 20057, USA

Cost of living in domed cities

SIR — Perhaps the most important result of the first Biosphere-2 experiment was the documentation of the extremely high cost of providing nature's free life support services with the use of non-renewable fossil fuels¹. Now that data on the energy consumption during the two-year closure (1991–93) have been published, the monetary cost can be calculated.

Electricity required to maintain a habitable environment artificially for two years was consumed at a rate of 700 kW and natural gas at the rate of 23 million kJ per hour (ref. 2). In units that utility companies use to calculate our bills, consumption for the two-year period comes to 12.5 million kWh of electricity and 3.8 million therms of gas. At 10 cents per kWh and 70 cents per therm (average rates in the United States), the costs would be \$1.25 million and \$2.6 million respectively, or approaching \$4 million altogether. (The underwriters of the Biosphere-2 experiment did not actually have to pay this much in dollars because they had their own power plants and could produce energy at commercial rates.) If the crew of 8 people had to pay the utility bills at these residen-

tial rates, their monthly bill would be more than \$150,000. At anywhere near this cost, very few of the billions of people on Earth could ever afford to live in domed cities.

Not all but a good portion of this energy cost can be considered as replacing the life-support services involved in the Earth's solar-driven atmospheric and hydrological cycles for which we pay little or no money. Most of the electricity, for example, was consumed in operating the complex pumping and filtering machinery (both inside the enclosure and in the outside accessory 'lungs' and cooling towers) that circulated the air, maintained its pressure and recycled and cleaned all the water. Much of the natural gas was expended in heating and cooling the living space. The cost of electricity and gas is, of course, only the tip of the iceberg of the millions of dollars that have already been spent in building the complex and will be spent maintaining the complex engineering in the future.

Some scientists have been critical of the first Biosphere-2 experiment as not being "real science" because the crew were not scientists, but people trained and selected for their ability to work together, grow all their own food, monitor and manage the plant and animal life, record oxygen levels and other vital necessities, man the pumps and engineering control devices, and especially for their willingness to live at a subsistence level for two years. Very few specialized scientists could qualify for such a job.

In a recent report³, the crew estimated that they spent about 45% of their waking hours growing and preparing food, 25% in maintenance and repair, 20% in communication (inside and with the outside) and 5% in carrying out small research projects, which left little time for relaxation and recreation. I happened to be present soon after the biospherians emerged from their two-year isolation and were being interviewed by the press. Someone enquired about their personal lives while inside and I heard one of the crew exclaim: "Look, we were so busy just trying to survive that we didn't have any time for hanky-panky". As with the Apollo-13 flight to the Moon, survival becomes the mission when life-support is in question.

The fact that the eight people emerged still speaking to each other and healthier than when they went in is, in itself, quite an accomplishment. Viewed as an exercise in human ecology, life-support science (biospherics) and environmental engineering, the experiment, in my opinion, was a considerable success.

Eugene P. Odum

Institute of Ecology,
University of Georgia,
Athens, Georgia 30602-2202, USA

1. Macilwain, C. *Nature* **380**, 275 (1996).
2. Dempster, W. F. *Society for Automotive Engineers* tech. pap. ser. 932290 (Warrendale, Pennsylvania, 1993).
3. Nelson, M. & Dempster, W. F. (unpublished results.)

Less rosy view of science funding

SIR — My view of science funding in the United States is neither as optimistic nor as complimentary to President Bill Clinton as *Nature's* (**381**, 1; 1996). Although Congress has generally supported funding of investigator-initiated academic research, the president repeatedly emphasizes the federal government's own "critical role to play" in carrying out and directing research and development. Moreover, he derides the budget-cutters when they find needed funding by trimming ineffective technology programmes at the Department of Commerce and the Environmental Protection Agency (EPA).

While small increases in Clinton's budget for fiscal year 1997 appear to keep the National Institutes of Health (NIH) and the National Science Foundation ahead of inflation, buried in the fine print are projections of sizeable reductions in extramural grants, and in many science and technology programmes at other agencies. For example, excluding the funding earmarked for refurbishing the NIH's on-campus hospital, its actual budget increase is only 1.3%, well below the rate of inflation.

The closer one looks, the more dire the implications. Clinton administration largesse is targeted to the agencies that historically have funded the lowest quality and most applied research. For example, some of the largest proposed increases are for the Department of Commerce (up 16.9%, to \$4.3 billion) and the EPA (up 22.8%, to \$7 billion). In 1995, the EPA was harshly criticized by a National Academy of Sciences panel for low scientific standards and lack of peer review of its research — problems that more money alone cannot solve.

Even without new funds, Clinton could easily adopt several no-cost strategies that would increase productivity. He could limit government support to highly meritorious projects that are unlikely to be undertaken by the private sector. He could craft incentives to increase private-sector participation in early stages of research. He could enhance the overall quality of government-sponsored research by subjecting all proposals to a scientific peer-review process to ensure quality and merit. But none of these seems to have occurred to the president or his advisers.

Henry I. Miller

Hoover Institution & Institute for
International Studies,
Stanford University,
Stanford, California 94305-6010, USA
e-mail: miller@hoover.stanford.edu

corres@nature.com

Letters submitted for Correspondence should be typed, double-spaced, on one side of the paper only, or e-mailed to corres@nature.com.

A biological marker for dyslexia

Chris Frith and Uta Frith

Brain imaging reveals that, in dyslexics, the visual-motion area of the brain fails to activate. This failure may be a marker for a deviation in brain function which remained largely invisible until complex writing systems evolved.

THERE are many reasons for a child failing to learn to read. He (for the problem is greater in boys) may not be exposed to print, he may be badly taught or he may hate school. However, about 5–10 per cent of children have difficulty learning to read in spite of the best possible circumstances¹. In these cases, the difficulty is believed to reflect neurological problems and the disorder is called developmental dyslexia. The problem tends to run in families: if one of the parents is dyslexic, then almost half of their children can be expected to have reading difficulties².

The difficulties experienced by these children are not restricted to reading: even before learning to read they have subtle problems with language³. The streams of sound which compose spoken language are infinitely varied, but are constructed from a small number of component phonemes (subsyllabic units). The distinctions between these phonemes are largely determined by the way in which they are articulated (for example, the position of the tongue in the mouth). In order to repeat aloud a word that we hear, particularly if it is a new word such as 'cernalit', we have to be able to segment the sound into its component phonemes.

Dyslexic children have small, but detectable, difficulties with word-repetition tasks even before they learn to read. Even adult developmental dyslexics who have eventually achieved a reasonable proficiency in reading still have problems if asked to repeat long words that are new to them⁴. In alphabetic scripts such as English, we have to learn how to map phonemes (rather than whole words) onto arbitrary graphic signs. A problem with splitting the sounds of words into their component phonemes is a major disadvantage for learning to read⁵.

The story so far provides a plausible account of developmental dyslexia at the cognitive level, but says nothing about the underlying physiological problem. On page 66 of this issue⁶, Guinevere Eden and her colleagues report a striking difference in the brain activity of adults with developmental dyslexia. These volunteers showed typical phonological problems. But the physiological anomaly, revealed by functional magnetic resonance imaging (fMRI), was elicited by viewing moving dots.

A fundamental principle of brain or-

ganization is functional segregation, which is currently best exemplified in the visual system. Different attributes of the visual scene, such as form, colour and motion, are routed to different regions of the extrastriate cortex. The region that



Correlates of brain activity and dyslexia — here, activity in the left hemisphere (top, of 'normal' brain; bottom, of dyslexic brain) associated with a simple rhyming task⁶. The task involved the visual presentation of two letters, and subjects had to indicate whether they rhymed (B–T) or not (B–K). To perform the task, one needs internally to evoke the name of the letter (/bee/, /tee/) and separate out the beginning phoneme from the rhyming part (/ee/). In controls, three areas were activated — Broca's area, Wernicke's area and the insula, a part of the cortex that lies between the various language areas; in dyslexics, only one — Broca's area — was activated. It can be assumed that in more challenging language tasks this reduced amount of activation would impair performance. Eden *et al.*⁶ now demonstrate that dyslexics also show reduced brain activation when carrying out a non-language test: the visual area V5 is not activated when moving dots are observed.

responds most strongly to visual motion is labelled V5 and is located in a lateral area of the prestriate cortex at the junction of the occipital and temporal lobes⁷. The dyslexic volunteers in Eden and colleagues' study differed strikingly from controls in that viewing randomly moving

dots failed to activate this area (see Fig. 2 of the paper, page 69). Along with this physiological difference, the dyslexic volunteers had a slight, but significant, impairment in detecting visual motion.

Eden *et al.* present these observations as further evidence for a selective deficit in the magnocellular subsystem of the brain in dyslexics. In the visual system, neurons in the magnocellular layers of the lateral geniculate nucleus (LGN) have rapid responses and are sensitive to motion, whereas neurons in the parvocellular layer have slow responses and are sensitive to form and colour. The input to V5 is dominated by the magnocellular stream. Direct evidence of abnormalities in the magnocellular layers of the LGN has been found in post-mortem studies of developmental dyslexics⁸. However, abnormalities are also observed in the language areas of the left temporal lobe in both post-mortem and functional imaging studies (see figure, left)⁹.

So we have two different strands of evidence — first, that dyslexics have a language problem specifically associated with the understanding of phonology; second, that there are behavioural and physiological abnormalities in the processing of visual movement. How do we reconcile these two strands, and what has the latter to do with reading?

There has been much speculation about visual problems of one kind and another associated with dyslexia. But the precise nature of these problems is not well defined and ranges from eye movement abnormalities to 'visual stress'. It is not clear whether there is a subgroup of dyslexics who primarily have some visual anomaly or whether the visual problem is associated with

the phonological problems seen in most cases.

In the volunteers studied by Eden *et al.*, phonological and visual problems seem to coexist. A defect in the processing of visual motion has been described before. However, the precise nature of the deficit shown by dyslexics is such that it is very unlikely to be a direct cause of reading difficulties, particularly as it does not affect vision under normal lighting conditions¹⁰.

What, then, are the implications of this result for our understanding of the physiological basis of developmental dyslexia? One possibility is that the lack of activity of V5 is simply a marker of a genetic deviation which manifests itself in a number of ways that are unrelated at the cognitive level; it could be that it is another of these manifestations that leads directly to reading difficulties.

A more interesting possibility is that the defect in the perception of visual motion is a marker of a more general cognitive deficit in timing which affects all brain modalities. For instance, in order to hear the differences between consonants, we need to be able to distinguish between very rapid changes in frequency. A subtle deficit in sensitivity to rapid changes in the auditory system could lead to a problem in the distinction between different phonemes¹¹, which would lead in turn to problems in learning to read. As yet, we know very little about functional segregation in the auditory system. The new brain-imaging techniques should enable us to identify those brain areas which are specialized for the representation of rapid auditory changes.

If the result reported by Eden and her colleagues can be replicated, and is shown to be specific to dyslexia, then we have a biological marker for this disorder which is independent of reading. Furthermore, V5 becomes fully myelinated before birth⁷, which suggests that its connections are already matured and unlikely to change. Thus, early detection of a physiological anomaly might be possible. The fMRI technique is an entirely non-hazardous procedure which could also be used to study how appropriate brain systems develop as a child learns to read.

Finally, what of the evolutionary and

cultural aspects of this topic? Written language has existed for a short time (5,000 years or so), and until recently has been used by only a small proportion of people. In preliterate cultures, the processing anomalies associated with dyslexia remain largely invisible: it is unlikely that special brain systems have evolved to handle written language. Future brain-imaging studies of dyslexic and normal readers will

provide a unique view of the different ways in which brain development interacts with culture. □

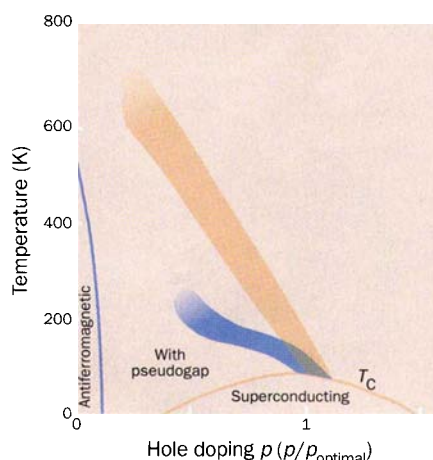
Chris Frith is in The Wellcome Department of Cognitive Neurology, Institute of Neurology, Queen Square, London WC1N 3BG, UK. Uta Frith is in The MRC Cognitive Development Unit, 4 Tavistock Street, London WC1H 0BT, UK.

SUPERCONDUCTIVITY

Crossovers in cuprates

B. Batlogg and V. J. Emery

FEW things in condensed-matter physics are more striking than states with long-range order. Crystals, magnets, superfluids and superconductors all derive their peculiar properties from the cooperative behaviour of the system as a whole. They all start out with a sharp phase transition, but in other situations the range of the order is not so long, and the transition is not so sharp — there is simply a gradual crossover in behaviour as the system cools down or is modified in some other way. Crossovers



The temperature-doping phase diagram of a cuprate superconductor. The shaded areas are crossover regions, where a property of the material changes gradually. In the upper crossover region, local antiferromagnetic correlations develop. The lower crossover region marks the area where the material begins to display a pseudogap in its electronic energy spectrum, suggesting the onset of local hole pairing. This region is not so well characterized experimentally, but it seems that the two crossovers are distinct at lower doping and merge close to optimal doping.

have much to teach us about the nature of these systems, but they are not easy to diagnose, especially if several things are going on at the same time. So it is with the cuprate high-temperature superconductors¹ (HTS) which, for some time, have given indications of having one or more crossovers that anticipate the true superconducting order, while leaving room for argument about what exactly is happening.

Angle-resolved photoemission experiments reported by Loeser *et al.*², Marshall *et al.*³ and Ding *et al.* (page 51 of this issue⁴) give us more specific information on the onset of superconducting correlations.

Copper oxide superconductors crystallize in layered structures, with stacks of a few CuO₂ sheets alternating with 'charge reservoir' layers. In (La, Ca)₂CuO₄, the prototypical cuprate superconductor discovered ten years ago by Georg Bednorz and Alex Müller¹, these building blocks are LaO₂ layers. Partial substitution of trivalent lanthanum by divalent calcium transfers electrons out of, or equivalently introduces holes into, the CuO₂ layers. By now, creative crystal chemists have synthesized dozens of different cuprate superconductors with a variety of charge reservoir layers. Through compositional variations of these layers, the hole doping level p can be adjusted, which has led in the past few years to a major advance in this field. A great commonality among all the different cuprates has been recognized: the essential physical properties, such as charge and spin dynamics, vary in a systematic way with p (see figure). For instance, the cuprates are insulating antiferromagnets for $p = 0$ and become superconducting metals with the relatively highest T_c at the optimal hole concentration of about 0.20 ± 0.05 per CuO₂ unit.

Present experiments suggest that there are two crossovers in these materials, which merge at higher levels of doping (see figure). When the system is cooled down below the upper crossover temperature, the CuO₂ planes develop local (short-range) antiferromagnetic correlations which, for very low doping, ultimately evolve into global (long-range) magnetic order. At all but the lowest levels of doping, there is a second, lower, crossover temperature below which, according to the new photoemission studies, the system develops a 'pseudogap', that is, a suppression of the density of low-energy excited states. As the doping increases, the two crossovers come together and finally merge at a point close to the true superconducting transition.

Evidence of a pseudogap has been seen

- Shaywitz, S. E. *et al.* *J. Am. med. Ass.* **264**, 998–1002 (1990).
- Pennington, B. F. *J. Child Neurol.* **10** (Suppl.), S69–S76 (1994).
- Scarborough, H. *Child Dev.* **61**, 1728–1743 (1990).
- Apthorp, H. S. *J. Learn. Dis.* **28**, 342–352 (1995).
- Bradley, L. & Bryant, P. *Nature* **271**, 746–747 (1978).
- Eden, G. F. *et al.* *Nature* **382**, 66–69 (1996).
- Zeki, S. *A Vision of the Brain* (Blackwell Scientific, Oxford, 1993).
- Livingstone, M. S. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **88**, 7943–7947 (1991).
- Paulesu, E. *et al.* *Brain* **119**, 143–157 (1996).
- Cornelissen, P. *et al.* *Vision Res.* **35**, 1483–1494 (1995).
- Tallal, P. *Behav. Brain Sci.* **6**, 219–220 (1983).

for a long time in a variety of experiments that probe the thermodynamics of charge and spin excitations. The contribution of the new experiments is to show that the pseudogap is also present in single-particle excitation spectra, and that it has essentially the same magnitude (25 to 35 meV, or equivalently 300 to 400 K), and dependence on direction relative to the Cu–O bond, as in the state below T_c . The suggestion is that below the second crossover temperature the CuO₂ planes develop local superconducting correlations which ultimately (at temperature T_c) evolve into global superconducting order.

In some respects, this behaviour is reminiscent of granular superconductors, where pairing is established within the grains at a crossover temperature T_{GR} , but true superconductivity does not appear until a much lower temperature T_c because of phase fluctuations allowed by the poor communication between the grains. In doped Mott insulators, such as the HTS, phase fluctuations are enhanced because only a small fraction of the available charge carriers participate in the superconductivity⁵. This observation is independent of the microscopic physics of the HTS.

Tunnelling experiments in granular lead⁶, for example, show that T_{GR} is roughly equal to the transition temperature of the crystalline material, and is about half the zero-temperature energy gap, a hallmark of the outstandingly successful theory of superconductivity proposed by Bardeen, Cooper and Schrieffer⁷ in 1957. As shown in the figure, the lower crossover in the cuprates shows a similar behaviour. Further studies are required to explore the similarity between granular superconductors and the HTS, and to improve our understanding of the low-doping region. In particular, it is not clear to us which crossover the low-doping point in Fig. 4 of Ding *et al.* belongs to.

The figure leaves little doubt that something new is going on which involves, at the very least, a connection between antiferromagnetism and superconductivity. Opinions differ about the precise nature of this connection, however, and about the fundamental physics of the HTS. □

B. Batlogg is at the Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974, USA. V. J. Emery is at the Brookhaven National Laboratory, Upton, New York 11973, USA.

Low-down on a land bridge

Paul Colinvaux

THE land of Beringia, as the old Bering land bridge has come to be called, exists whenever the shallow Chukchi and Bering Seas to the north and south of Bering Strait are drained by the fall in sea level — 100 m or more — that comes with every ice age. Beringia then serves as a land bridge joining North America to Asia by a Siberian connection more than 1,500 km wide.

But what was Beringia's environment like during the last glaciation, particularly its maximum 22,000 to 14,000 years ago, when much more land was exposed than at present? Elias *et al.* (page 60 of this issue¹) provide some compelling evidence from under-sea cores that swings a long-standing debate firmly one way. That debate is over whether Beringia largely consisted of tundra (treeless Arctic vegetation on frozen ground, with sedges more prominent than grasses), or steppe (dry grasslands like those of Central Asia or the colder North American prairies), and thus of what kinds of life the land could have supported. Steppes, like prairies, typically support grazing animals and their predators. Tundras, the lands of lemmings and caribou, are less productive. Humans might well find more to eat on a steppe than a tundra.

Erik Hultén gave us the first clue to Beringian vegetation 60 years ago, when he showed that the forests of Siberia and Alaska were so different that they could not have merged in the Pleistocene² — the epoch of ice ages, roughly the past two million years. Beringia must have been treeless. Next came pollen analysis of lake mud as adjacent regions, particularly in Alaska, were searched for ancient land-bridge lakes. The pollen data confirmed Hultén's insights: there is no evidence of trees in glacial-age deposits. But all the early Alaskan records had a peculiarity in that they grouped the pollen of grass, sedge, heath, willow and tundra shrubs, with extraordinary amounts of pollen of sage (*Artemisia*). No modern tundras yield pollen rich in sage.

A treeless land with *Artemisia* has an almost irresistible connotation of steppe — if not the actual sagebrush prairie of North America, then the steppe-tundra described by Russian workers in northern Central Asia. The grasses of steppes typically support herds of game, thus allowing the hypothesis of Beringia as a vast grassy plain supporting grazing animals. Bones of mammoths, horses,

bison and other ungulates are unearthed in great numbers in the remnants of Beringia, particularly in Alaskan gold diggings, and dates on the bones now virtually blanket the part of the glacial interval accessible to radiocarbon dating. These, then, could have been the animals of the postulated 'mammoth-steppe' of Beringia. Human hunters could have preyed on these animals, eventually debouching from eastern Beringia into North America.

Most of the pollen analysts of Alaska have resisted this hypothesis, claiming that



Beringian lowlands late in the glacial cycle — as represented by the uplands of the modern Seward Peninsula, Alaska.

their pollen suggested a bleak tundra, despite the sage, or even actual polar desert. But what we argued over was in fact land now beneath the sea, where we had no records because our Alaskan lakes were of necessity on Beringian highlands. Elias *et al.* at last provide records from the flooded land surface of old Beringia itself, which should go far to settle the matter. Beringia was tundra, not steppe.

A programme of ship-borne coring recovered, in 20 cores, peat from mires lying under marine sediments at the bottom of the Bering and Chukchi Seas (see map, page 61). Collectively, the peats span from early in glacial times to the final flooding of the bridge, dated by the peat at sill depth to shortly after 11,000 years before the present. Pollen assemblages at all 20 sites are without significant *Artemisia* pollen. Instead, pollen from these Beringian bottom peats matches pollen from modern mesic tundras of Arctic Alaska (see photograph): these are tussocky 'fields' with prostrate willows, tiny (dwarf) birches that are often shorter than the sedge tussocks between which they grow, and a characteristic mass of spring flowers. Plant macrofossils in the peats are

1. Bednorz, J. G. & Müller, K. A. Z. Phys. B64, 189–193 (1986).

2. Loeser, A. G., Shen, Z.-X. & Dessau, D. S. Physica C263, 208–213 (1996).

3. Marshall, D. S. *et al.* Phys. Rev. Lett. 76, 4841–4844 (1996).

4. Ding, H. *et al.* Nature 382, 51–54 (1996).

5. Emery, V. J. & Kivelson, S. A. Nature 374, 434–437 (1995).

6. Barber, R. P. Jr *et al.* Phys. Rev. 49, 3409–3412 (1994).

7. Bardeen, J., Cooper, L. & Schrieffer, J. R. Phys. Rev. 108, 1175–1204 (1957).

1. Elias, S. A. *et al.* Nature 382, 60–63 (1996).

2. Hultén, E. *Outline of the History of Arctic and Boreal Biota During the Quaternary Period* (Bokförlägs Aktiebolaget Thule, Stockholm, 1937).

3. Roosevelt, A. C. *et al.* Science 272, 373–384 (1996).

of the familiar assemblies of tundra ponds.

Thus, the Beringian plains were covered with tundra, not steppe. Must we pollen analysts hang our heads? I do not think so. The sage pollen in our upland samples is real and represents what we have always said it did — a cold and dry tundra landscape on the windswept hills, poorly productive, in which frost boils and bare patches gave ample opportunity for several species of Arctic sage to flourish. Lowlands were mesic tundra; uplands were tundra sufficiently dry and impoverished for Arctic sage.

But what of the mammal bones? Ecological common sense makes it most unlikely that dense herds of big-game animals subsisted on a tundra. The evidence for tundra probably cannot be made to go away with the suggestion that the bog records were too local: steppe pollen blows, the mires were sampled at 20 places over a huge area, and mires are classic traps for wind-blown pollen. Moreover, the study includes extra evidence in the form of beetle remains identified as species which, like the pollen, deny the existence of steppe. So there was no steppe, only tundra — with mammal bones. The simplest way out of the dilemma is the one offered by pollen analysts in the past, which is that the bone assemblies represent rare individuals or populations wandering far from optimum habitats in response to crowding or runs of good years. Doubtless, a spirited response from the big-mammal community will be forthcoming.

These new data confirm that the people who spread from Beringia to North America about 14,000 years ago came from stock able to gain a livelihood from the meagre resources of tundra, or perhaps from the sea coast. They must have been capable people, skilled in making Arctic clothing, among other traits. If, as we think, it was their descendants who first populated North America before crossing a second land bridge in Panama to reach South America, they would have needed to have been adaptable as well as capable.

Direct evidence of the climate and vegetation of the Panama land bridge in the late glacial times of human passage is less complete than that for Beringia, although it is known that both temperature and precipitation changed. Some parts were arid, though evidence persuasive to some suggests that there were tropical forests to be crossed on the line of the modern canal. If people lived in the forested Amazon lowlands as early as 11,000 years ago, as recent studies suggest³, the acquisition of suitable habits might have begun on the Panama land bridge itself, by people whose adaptability had once fitted them for life in Beringia. □

Paul Colinvaux is at the Smithsonian Tropical Research Institute, PO Box 2072, Balboa, Republic of Panama.

A solar system next door

Gordon A. H. Walker

THE chances of directly imaging a system of planets around another star have improved dramatically. At the June meeting of the American Astronomical Society*, George Gatewood of the Physics and Astronomy and the Geological and Planetary Science departments, University of Pittsburgh, announced the detection of a highly significant acceleration in the nearby star Lalande 21185. This star's motion shows perturbations similar in size to those induced on the Sun by the planets of the Solar System. He tentatively attributes the acceleration to as many as three planets with periods of several years, which would make this by far the most similar planetary

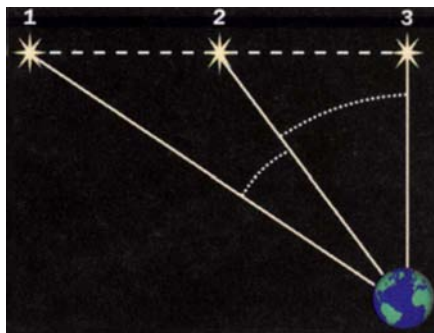


FIG. 1 The relative motion of a nearby star and the Sun generates perspective acceleration. The star is observed three times as it moves along the dashed line, moving equal distances between the observations while the angular motion seen from the Earth appears to increase.

system to our own yet discovered, although these attributions are less confident than the overall acceleration itself.

Despite being only 8 light years from the Sun, Lalande 21185 is too faint to be seen with the naked eye, but it is an easy object for a small telescope. It is cooler than the Sun (3,200 K at the surface, compared with 6,200 K) and only about 40 per cent as massive, but this tiny star has now become the prime candidate for a direct view of another planetary system.

Interest in extrasolar planets among astronomers and the general public has been intense over the past six months, following a series of discoveries of massive planets around more distant stars. The National Aeronautics and Space Administration, for instance, has made the detection of an Earth-like planet beyond our Solar System a keystone of its Origins Program. But we still do not know how common systems of planets might be, or how closely they resemble the planets of our Solar System.

Until now, surveys have looked for

the very subtle, long-term acceleration induced by the mass of an unseen planet either by carefully monitoring a star's radial velocity or by astrometry — tracking its motion on the sky. As we are probably looking for planets less than one-thousandth the mass of their primary star, these observations are extremely exacting. But efforts appeared finally to pay off in late 1995 and early 1996 with detections of Jupiter-mass companions to several Sun-like stars (currently at least five) from precise radial-velocity measurements^{1,2}.

Unfortunately, most of the candidates lie much closer to their parent star than the Earth does to the Sun. Those small separations, and the fact that the stars lie more than 10 parsecs (33 light years) from Earth, mean that they will be hard to detect directly. At that distance, the Earth's orbital radius would subtend an angle of only 0.1 arc seconds, so it would be impossible to distinguish a planet in an even smaller orbit from the much brighter star in a direct image using current technology. But that is not so for the planets proposed for Lalande 21185.

An astrometrists measures the very small displacements of a star's image against those of background stars. From 50 years of photographic observations³, and multi-channel photoelectric observations for the past 8 years with the Allegheny Observatory 0.76-m Thaw refractor, Gatewood and his colleagues have carefully determined the acceleration of Lalande 21185. At 2.6 parsecs away, Lalande 21185 is the sixth closest star to us, and for a star so close to the Sun, relative motion produces a so-called perspective acceleration (Fig. 1) which also had to be subtracted. Unlike the spurious detection in 1969 of a system of planets around Barnard's Star which was also based on photographic astrometry, there seems little doubt about the acceleration of Lalande 21185.

The results from the photographic data are shown in Fig. 2, after correction for the effects of parallax and proper motion. Most of the curvature is caused by perspective acceleration, which can be accurately predicted from the distance and motion of the star relative to the Sun, both of which are well known for Lalande 21185. At Allegheny they find that the predicted and observed accelerations differ by 6×10^{-5} arc seconds per year per year, which is 7.5 times the standard error of the observations. Gatewood has calculated that this departure can be explained by the presence of up to three Jupiter-mass planets, with periods of 5.8, about 30 and more than 30 years, but these are only significant at the two- to three-sigma level. A binary companion star can be ruled out because

*188th Meeting of the American Astronomical Society, Madison, Wisconsin, USA, 9–13 June 1996.

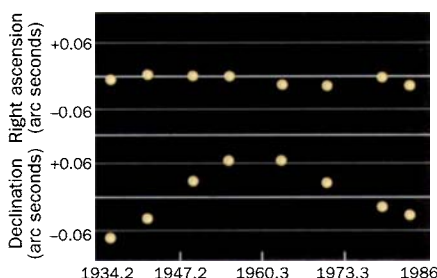


FIG. 2 The motion of Lalande 21185 from 50 years of photographic data taken at the Allegheny Observatory, after correction for the effects of parallax and proper motion. The points are 6-year means. Although most of the curvature is caused by perspective acceleration, it departs significantly from that predicted. (Figure kindly supplied by G. Gatewood.)

it would have been seen long ago.

The innermost companion, for which the orbit appears to be circular, would have a maximum separation of about 1 arc second from the primary. That makes direct imaging a real possibility, despite the glare from the star, using a large ground-based telescope (with a mirror greater than 3 metres in diameter) at a good site, or the Hubble Space Telescope. Only by capturing an image of such an object and eventually securing a spectrogram can we be sure it is a planet.

If it shines largely by reflected light, the image of the 5.8-year planet may only be some 10^{-9} as bright as the star. But with the introduction of adaptive optics and the refurbishment of the Hubble Space Telescope, the quality of images that can now be achieved both from the ground and from space at near-infrared wavelengths (1 to $2\ \mu\text{m}$) has recently improved enough to make detection feasible.

Diffraction sets the ultimate limit to angular resolution, but, on the ground, image-spread is dominated by rapidly changing thermal turbulence in the atmosphere. Thanks to adaptive optics, however, wavefront distortions can be sensed and corrected in real time with a deformable mirror. This, in principle, makes near-diffraction-limited imaging possible, at least at near-infrared wavelengths. Several such adaptive optical systems of differing complexity are now in use on large ground-based telescopes, and some have coronagraphic optics which exclude the great majority of contaminating light from the star. So the race is on to obtain a convincing image of a planet in the Lalande 21185 system, a race that might be won as early as next year. □

Gordon A. H. Walker is in the Physics and Astronomy Department, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4.

Changing times, changing teams

Henry Gee

LAST Wednesday (26 June), England was beaten by Germany in the semi-finals of the European soccer championships, Euro '96. The match was settled by a penalty shoot-out, and was an uncanny replay of the 1990 World Cup — a semi-final win for Germany over England, again decided by penalties. But the parallels run deeper still: through a World Cup quarter-final meeting in 1970 (Germany the winners), and back to the 1966 World Cup final, with England this time the victor.

But how real are the comparisons? Nominally, the teams represented the same adversaries, but the individual players were different. Of those on the pitch last Wednesday, only five played in 1990, and, of course, there were none from 1970 and 1966. Indeed, the England captain in 1966, Bobby Moore, died recently, and most of the current squads weren't even born then. This is to be expected, as national teams carry on, eclipsing individual careers.

If soccer were as we might assume ecology to be, Bobby Moore would have led the England team last Wednesday: it is easy to think that the faunal associations we see today are as they have ever been, the individual members bound together in equilibrium by links of adaptation and competition. Communities, one assumes, respond to factors such as climate change by wholesale shifts. For example, cooling would drive deciduous forests into southern refugia, and their animal residents, the whole team of them, with them.

Palaeontologists have long known that history, like soccer, tells a different story, but this has probably never been documented in the exhaustive detail that Graham *et al.* have achieved in their study of the faunal history of the United States over the past 20,000 years (*Science* **272**, 1601–1606; 1996). They have collated species lists of mammal faunas from 2,945 localities in the contiguous United States, and have developed a software package called FAUNMAP which displays this information in space and time.

Individual species can be picked out, and their ranges followed as the climate changes. Associations between species can also be measured, using modern examples as standards. This has allowed researchers to find out whether modern-day ecological associations have long and distinct histories, or have been thrown together more recently. FAUNMAP can also be used to measure spatial heterogeneity; that is, the extent to which faunal associations differ with distance.

FAUNMAP analyses show that species are not team players — they tend to look after themselves, not their associations, and respond to climate change in idiosyn-

cratic ways. At the end of the Pleistocene, 10,000–15,000 years ago, some species moved northwards, but at different rates. Others, though, moved eastwards, while the ranges of some western species contracted. Still others have stayed put, their present distribution being the same as it was during the coldest parts of the ice age 15,000–20,000 years ago.

The implications of this work are twofold. First, the uncoordinated movement of species in response to climate change produces novel combinations, makeshift alliances rather than associations forged by co-evolution and competition over millions of years. Second, the ecological stability we see today is illusory — modern communities emerged only in the past few thousand years.

Answers to larger questions of biogeography and spatial heterogeneity are sensitive to scale. At the largest scales, Graham *et al.* show that present-day biogeographical 'provinces' have deep roots, even though many of the participants have changed. The faunas of the United States appear to be divided into eastern and western components along the 100-degree line of longitude, whether one looks at Late Pleistocene or Holocene faunas (respectively before and since the conventional 10,000-year watershed). This division marks a boundary between the forests and grasslands to the east, and the deserts and mountains to the west.

Closer inspection reveals significant differences. For example, the Late Pleistocene faunas of the eastern United States are divided into distinct northern and southern components, presumably dictated by temperature. Even so, within any given province, one can use the roster of species to discern environmental preferences, even if the membership of each list differs from place to place. There'll always be English and German soccer teams (at least, one hopes so). But the individual players will have changed.

Faunas from any two places share progressively fewer species with increasing distance between the localities, whether one is looking at Late Pleistocene or Holocene faunas. But for any given distance, Late Pleistocene faunas are less similar than Holocene faunas. This suggests that spatial heterogeneity in the Pleistocene was greater than today. Faunas responded to a more intricate patchwork of environments with a greater variety of species. The environment of the modern United States is more homogeneous, and there are fewer distinct faunal associations — in effect, fewer teams are up for the cup. □

Henry Gee is an assistant editor at Nature.

1. Mayor, M. & Queloz, D. *Nature* **378**, 355–359 (1995).
2. Marcy, G. W. & Butler, R. P. <http://cannon.sfsu.edu/~williams/planetsearch/planetsearch.html>
3. Gatewood, G. *Astr. J.* **94**, 213–224 (1987).

True grit spells double trouble

Scott Lehman

OCCASIONALLY, geophysicists bump into layers of sand on the flanks of the Mid-Atlantic Ridge in places far from any debris-carrying turbidity flows, and puzzle over their origin¹. But in 1988, when Hartmut Heinrich encountered similar layers on a seamount in the northeastern Atlantic, their significance became clear: rather than a stately procession of growth and decay, the continental ice caps that once surrounded the North Atlantic basin have undergone a series of massive convulsions, releasing throngs of debris-laden icebergs to the sea every 6,000 to 12,000 years in events lasting just a few hundred years or so². The layers have since been found over a broad swath of the North Atlantic between about 55° and 40° N, where bergs carried south from the ice sheets met the warm waters of the subtropical gyre, where they melted and deposited their debris³.

Despite the emergence of a veritable cottage industry devoted to the study of these events, the question of how and why they occur remains. Do they record a recurrent internal instability of the Laurentide ice sheet, or a response to some form of global climate forcing⁴? In two studies appearing back-to-back in *Paleoceanography*, Gwiazda *et al.*⁵ and Revel *et al.*⁶ use isotope geochemistry to take up the question once more. Interestingly, they appear to reach opposite conclusions.

Both studies seek to determine the geological source of ice-rafted detritus within Heinrich layers, in order to infer which ice caps underwent accelerated calving. It has already been shown beyond doubt that the Laurentide ice sheet of North America was a major player — Heinrich layers thicken towards Hudson Strait and Baffin Island⁷, and are tainted by grains of uniquely Canadian carbonate^{3,8}. Demonstrating that the Laurentide acted alone would reinforce early claims of glaciological modellers that the ice sheet is capable of unprovoked binge-purge behaviour⁹. If instead, other ice caps surrounding the North Atlantic basin were also involved, it might be that they responded in common to climatic forcing. The question is by no means strictly of geological interest. If the events were coerced, the associated forcing may still be operating, colouring the backdrop of natural climate variability against which we try to discern anthropogenic climate changes (as Nicholls discusses on page 27).

Gwiazda *et al.* concentrate their efforts on the study of the second to last of the Heinrich events, called H-2, which occurred just before the height of the last glaciation, about 21,000 years ago accord-

ing to radiocarbon dating. They measured the lead isotope composition ($^{207}\text{Pb}/^{204}\text{Pb}$ versus $^{206}\text{Pb}/^{204}\text{Pb}$) of 15 to 30 single feldspar grains, hand-picked from H-2 and from typical glacial sediments above and below, for a number of cores collected along the iceberg trajectories from different possible source areas (see figure). At each site, grains from the H-2 layer showed lead isotope trends consistent with their being derived almost exclusively from the Churchill Province, at the theoretical heart of binge-purge country.

A terrane similar in age and lead isotope composition to the Churchill Province occurs in the central portion of the Scandinavian shield, but is flanked seaward by much younger, relatively ^{206}Pb -rich terranes. As ice reaching the sea could not have incorporated materials from the central part of the shield without also picking up grains from the younger areas, the authors discount the Fennoscandian ice sheet as a source of ice-rafted grains during H-2. Of course, the likelihood of missing rare contributions is appreciable when analysing only a few tens of grains (a practical limit), so Gwiazda *et al.* also analysed composite samples containing 70 to 300 grains, and obtained results consistent with their interpretation of the single-grain data.

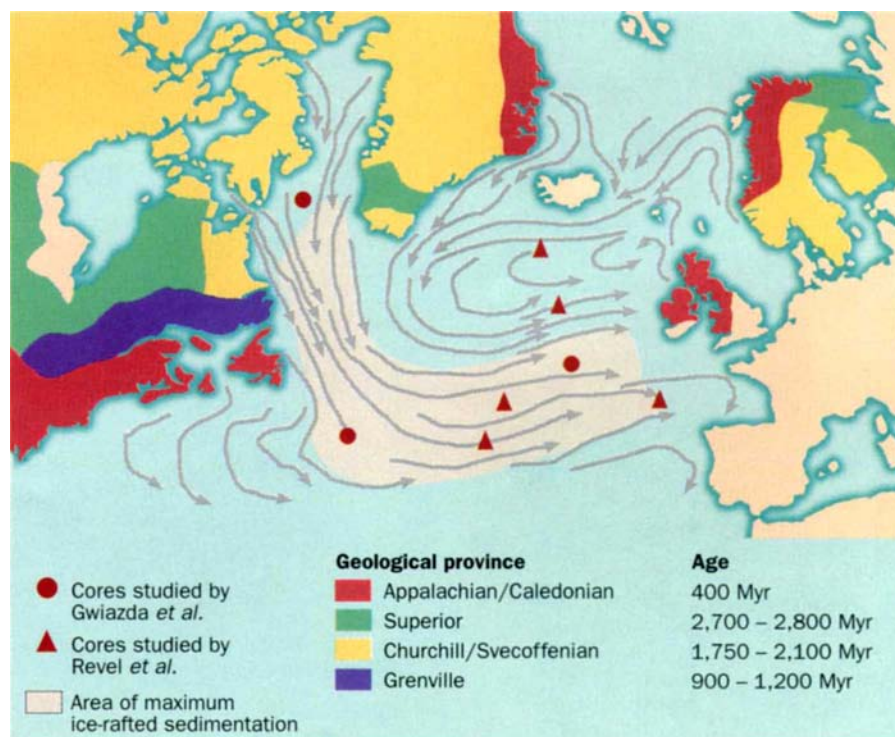
In contrast to the H-2 grains, single

grains from the sediments bounding H-2 were of varied origin, with source areas in the Churchill and Grenville Provinces, the Appalachian range, and possibly Scandinavia and Greenland, indicating a more geographically uniform contribution of iceberg detritus outside Heinrich events.

Revel *et al.* made measurements of strontium concentration and isotopic composition ($^{87}\text{Sr}/^{86}\text{Sr}$) in bulk samples, acid-leached to remove plankton shells which carry the strontium isotopic composition of sea water, from the last six Heinrich events and from surrounding glacial and interglacial sediments in cores not far from some of those studied by Gwiazda *et al.* Although it is difficult to differentiate between the North American and Scandinavian shields on the basis of strontium isotopic composition alone, Revel *et al.* argue that the consideration of strontium concentration data makes such distinctions feasible.

Despite my own concerns over spatial variability of strontium concentrations, the results display geographical patterns that make them believable. Cores from the central part of the belt of maximum ice-rafting show an overwhelmingly Canadian strontium signature during Heinrich events. At the eastern end of the belt, the signature veers towards that of Scandinavia, and possibly Britain. Just north of the belt it is increasingly British, and farther north, Scandinavian.

On the face of it, the two studies appear to be in conflict, but procedural differences leave some room for reconciliation. Gwiazda *et al.* analysed large (> 150 μm) feldspar grains from the swath



Location of the sediment cores analysed by Gwiazda *et al.*⁵ and Revel *et al.*⁶, with respect to the average trajectory of debris-carrying icebergs during glacial periods (arrows, after ref. 10) and possible geological source areas.

of maximum ice-rafting (originally mapped on the basis of sand larger than 200 μm ; ref. 10), whereas Revel *et al.* used bulk samples containing all grain sizes, and, unlike sand, small grains may be transported in suspension. It may be that icebergs also calved faster from Scandinavia, Great Britain and Iceland during Heinrich events, but tended to melt not far from port, where large grains were deposited. Fine-grained detritus may then have remained in suspension and been carried to sites in the open Atlantic. Such a picture is consistent with a variety of recent studies in which grains were characterized according to their optical properties^{4,11,12}.

From a methodological point of view, the study by Gwiazda *et al.* is clearly top-notch, and when they consider all the possible implications of their data, they concede that Heinrich events may have been driven by global climate variation, with the Laurentide showing greater sensitivity than the other ice sheets and thereby dominating the sedimentary signature. Recently updated glaciological models with more realistic boundary conditions¹³ imply just that: they seem to require a bit of a nudge in order to elicit binge-purge behaviour. Evidence of such forcing may be at hand from studies of the new ice cores from the Summit site on Greenland (GISP2 and GRIP) which, like an increasing number of palaeoclimate records from around the world, appear to bear the stamp of Heinrich events¹⁴.

In an analysis to appear in a special issue of the *Journal of Geophysical Research*, Paul Mayewski and colleagues¹⁵ show that the annually-dated climate record of the past 110,000 years in the GISP2 core can be decomposed into a set of statistically significant cyclic components. One of these, the 6,100-year cycle, is associated in time with each of the Heinrich events and with the Younger Dryas cold period. It also appears to

extend into the Holocene (the past 10,000 years) where it coincides with a mid-Holocene cold period and, later, the Little Ice Age¹⁶.

So the timing of Heinrich events may have been organized by a regular pattern of climate forcing that persists to this day. What that forcing might be is very difficult to say. Mayewski *et al.* suggest that it is a nonlinear response (a low-order harmonic) to the same orbital cycles that are thought to drive the waxing and waning of the ice sheets at lower frequencies.

Regardless of whether Gwiazda *et al.* and Revel *et al.* come to see eye to eye, they

may soon find themselves debating with an entirely different breed of scientist. Because, if the Earth is still riding the cyclic rebound from a relatively cold Little Ice Age, the greenhouse sceptics will inevitably claim that much of the 0.5 °C global warming of the past century would have occurred even without fossil fuel burning. But if mother nature and humankind are each conspiring to heat up the Earth, I'm not inclined to be so sanguine. □

Scott Lehman is in the Institute of Arctic and Alpine Research, and the Department of Geological Sciences, University of Colorado, Boulder, Colorado 80303, USA.

CLIMATE CHANGE

An incriminating fingerprint

Neville Nicholls

THERE is a great deal of evidence that global air temperatures have changed during the twentieth century. It has, however, proved difficult to identify the causes of these changes. Now, on page 39 of this issue¹, Santer *et al.* have provided the most convincing demonstration yet that human actions may have made a contribution. Their demonstration relies on comparing the expected climatic effects of changes in the concentrations of carbon dioxide, sulphate aerosols and stratospheric ozone, with the observed pattern of change in the atmosphere's temperature distribution.

The Intergovernmental Panel on Climate Change (IPCC) First Assessment Report², completed in August 1990, concentrated on the possible climatic effects of increases in greenhouse-gas concentrations, and concluded that climate warming should result from the increased atmospheric concentrations of these gases since the pre-industrial period. However, the report did not find clear evidence that the increase in greenhouse-gas concentrations had led, at that stage, to climate change.

The Second Assessment Report³, of which I was convening lead author, has just been published. It considers the possible effects of anthropogenic sulphate aerosols (which come from the combustion of fossil fuels and biomass burning) as well as greenhouse gases. The report concludes that "The balance of evidence suggests a discernible human influence on global climate". This conclusion emerged from pattern-based studies, in which the regional response of global climate models to recent changes in greenhouse-gas and sulphate-aerosol concentrations is compared with observed geographical patterns of atmospheric temperature change.

Pattern-based studies were used

because of the difficulties in determining the cause of the observed twentieth-century rise in globally averaged temperature. There are uncertainties in natural internal climate variability and in the histories of the many natural and human-induced external factors that affect the climate (volcanoes and increases in greenhouse-gas concentrations, for example), and many possible combinations of these factors could result in similar histories of global mean temperature. A way of getting around this problem is to look at the different geographical patterns of climate change likely to arise from the various relevant factors. The large number of degrees of freedom inherent in patterns of change, compared with those available from simply examining the global mean temperature, makes pattern-based methods more powerful tests for determining the cause of observed climate variations.

Several studies have compared patterns of near-surface temperature change with the changes expected from increases in greenhouse-gas concentration alone, and also with the combined effect of greenhouse gases and anthropogenic sulphate aerosols⁴⁻⁶. These studies produced similar conclusions, namely, that the observed patterns of temperature change were more similar to the predicted changes caused by greenhouse gases plus sulphate aerosols than to those from greenhouse gases alone. The pattern correspondence has increased over the past few decades, as would be expected from a human effect.

Santer and his colleagues¹ have now extended these studies by looking at vertical temperature patterns and also including stratospheric ozone depletion — being a greenhouse gas, this ozone loss tends to cool the lower stratosphere. The combined estimated effects of all three anthropogenic forcings were compared with

1. Kasameyer, P. W., von Herzen, R. P. & Simmons, G. *J. geophys. Res.* **77**, 3162–3167 (1972).
2. Heinrich, H. *Quat. Res.* **29**, 142–152 (1988).
3. Bond, G. *et al. Nature* **360**, 245–249 (1992).
4. Bond, G. B. & Lotti, R. *Science* **267**, 1005–1010 (1995).
5. Gwiazda, R. H., Hemming, S. R. & Broecker, W. S. *Paleoceanography* **11**, 77–93 (1996).
6. Revel, M. *et al. Paleoceanography* **11**, 95–113 (1996).
7. Dowdeswell, J. A., Maslin, M. A., Andrews, J. T. & McCave, I. N. *Geology* **16**, 999–1002 (1995).
8. Andrews, J. T. & Tedesco, K. *Geology* **20**, 1087–1090 (1992).
9. MacAyeal, D. *Paleoceanography* **8**, 775–784 (1993).
10. Ruddiman, W. F. *Geol. Soc. Am. Bull.* **88**, 1813–1827 (1977).
11. Rahman, A. *Geology* **23**, 487–490 (1995).
12. Fronval, T., Jansen, E., Bloemendal, J. & Johnsen, S. *Nature* **374**, 443–446 (1995).
13. Alley, R. B. *J. Glaciol.* **39**, 447–454 (1993).
14. Mayewski, P. A. *et al. Science* **263**, 1747–1751 (1994).
15. Mayewski, P. A. *et al. J. geophys. Res.* (in the press).
16. O'Brien, S. R. *et al. Science* **270**, 1962–1964 (1995).

vertical and latitudinal patterns of temperature change. The pattern seen in the model experiments^{4,7,8} used by Santer *et al.*¹ is one of tropospheric warming (greater in the mid-latitudes of the Southern Hemisphere), overlain with stratospheric cooling. A very similar pattern of change was found in the observations.

The study of Santer *et al.*, and those reported in the IPCC Second Assessment, show that an anthropogenic component of global climate change — the 'anthropogenic fingerprint' — may be appearing in the observed data. It must be pointed out, however, that this signal is the complicated pattern of change resulting from the combined effects of stratospheric ozone depletion and increased concentrations of greenhouse gases and sulphate aerosols. It does not mean that the effect of any one of these factors has been detected.

Many uncertainties remain in this work, and are acknowledged by Santer *et al.* and in the IPCC Second Assessment. There are uncertainties in estimates of the magnitude and patterns of the various natural and human-induced factors likely to affect climate. Climate models are far from perfect. There are also problems with the temperature observations: much of the similarity between the model and observed changes comes from the asymmetrical pattern of warming between the hemispheres, and the detection of that asymmetry relies on the relatively few observations of tropospheric temperatures in the Southern Hemisphere. Those observations are known to be affected by changes in instrumentation⁹, so further work is needed to eliminate the possibility that the observed asymmetry is the result of such changes¹⁰. In addition, the modelled hemispheric asymmetry depends on the way the models handle ocean heat uptake, and new studies¹¹ may lead to models with better representations of this process.

Despite these caveats, the results of Santer *et al.* — using the available data and state-of-the-art climate models — provide the clearest evidence yet that humans may have affected global climate. □

Neville Nicholls is in the Bureau of Meteorology Research Centre, PO Box 1289K, Melbourne, Victoria 3001, Australia.

Another role rolls in

Kim Nasmyth

MUTATIONS that affect the retinoblastoma tumour-suppressor protein (Rb), cyclinD/Cdk4 kinases and their p16 inhibitor are implicated in perturbed control of the cell cycle and in the genesis of a wide variety of human tumours¹. But why is Rb so important? An answer is emerging, as exemplified by the paper by White and colleagues² on page 88, the main element of which is that Rb regulates more aspects of the cell cycle than had been suspected.

Until recently, it was believed that Rb restrained cell proliferation in vertebrate cells by inhibiting factors, such as E2F, which are needed for the transcription of specific genes involved in promoting DNA replication³. An abrupt increase in Rb phosphorylation one or two hours before the onset of S phase of the cell cycle, brought about by cyclinD/Cdk4 protein kinase which accumulates at this stage, causes Rb to dissociate from E2F and allows it to activate S-phase genes^{4,5}.

So Rb is thought to impose a dependence of S-phase gene transcription on the prior activation, ultimately by growth factors, of cyclinD/Cdk4 kinases. Deregulated expression of S-phase genes, due either to hyperactivation of cyclin D or to loss of Rb, allows cells to enter S phase under circumstances that would normally not permit this, at least so the theory goes. But this idea never explained how loss of Rb enables tumour cells to increase their overall biosynthesis, without which an increase in the frequency of chromosome duplication (due to deregulation of E2F) would be lethal.

Last year, Cavanaugh *et al.*⁶ came up with data implying that Rb also regulates ribosomal RNA transcription mediated by Pol I, which provided hints as to how Rb might influence cellular biosynthesis as well as gene expression in S phase. They found accumulation of Rb in the nucleolus when monocyte-like cells were induced to differentiate *in vitro*; and they also showed that Rb can repress rRNA transcription *in vitro*, possibly by interacting with the upstream factor UBF. This paper did not demonstrate whether Rb really is involved in regulating Pol I *in vivo*, but it did raise the possibility that Rb might be a more general regulator of transcription needed for cell proliferation.

This notion is now significantly advanced by White *et al.*², who show that Rb represses most, if not all, genes transcribed by Pol III. Their paper demonstrates not only that Rb represses Pol-III-mediated transcription *in vitro*, but also that it performs this function in cultured primary mouse fibroblasts.

So, does this new finding shed light on the regulation of Pol III genes? At present,

not much. It will be important to establish how repression of Pol III transcription by Rb is regulated during the cell cycle and during differentiation. Whether Rb hyperphosphorylation in late G1 causes an increase in Pol-III-mediated transcription, like that observed for E2F-regulated Pol II genes, is not known. It will be interesting to know, furthermore, whether Rb imposes any dependence of Pol-III-mediated transcription on cyclinD/Cdk4 kinases.

As I have said, the importance of the Rb pathway in suppressing aberrant cell proliferation did not fully make sense while it was thought that it exclusively regulates S-phase-related genes through E2F. To appreciate this, we must remember that sustained cell proliferation depends on the duplication of all the constituents of the cell, not just chromosomes. With the exception of M phase, synthesis of metabolic enzymes, membrane compartments, the translational apparatus and so on occurs fairly continuously during the cell cycle⁷, whereas chromosomes are duplicated and segregated during discrete phases. To maintain a constant cell size, cells must ensure that chromosome duplication occurs no more frequently than duplication of the other cell constituents.

In microorganisms such as yeast or *Escherichia coli*, coordination between the chromosome cycle and biosynthesis of non-chromosomal constituents is largely achieved by a dependence of the chromosomal cycle on biosynthesis, and not vice versa. Key transitions within the chromosomal cycle (G1/S, G2/M or M/A) depend on cells growing to a critical size. Specifically blocking the duplication or segregation of chromosomes has little immediate effect on the rate of cellular biosynthesis. Furthermore, accelerating either of these processes, by hyperactivation of cyclin-dependent kinases that promote S or M phase in yeast, does not cause cells to grow faster (it merely causes them to divide with a smaller cell size)^{8,9}. DNA replication is, of course, ultimately necessary to sustain biosynthesis, but the key point is that DNA templates for transcription do not seem to be rate-limiting for biosynthesis in proliferating cells.

The relationship between cellular biosynthesis and the chromosome cycle is less clear in animal cells¹⁰. It seems, nevertheless, that growth to a critical cell size is important for entry into S phase¹¹, that blocking the chromosome cycle does not immediately block cellular biosynthesis, and that hyperactivation of S-phase-promoting kinases such as cyclinE/Cdk2 does not cause cells to divide more rapidly in the long term^{12,13}. Thus, even in animal cells, cellular biosynthesis might not be

- Santer, B. D. *et al.* *Nature* **382**, 39–46 (1996).
- Houghton, J. T. *et al.* (eds) *Climate Change: the IPCC Scientific Assessment* (Cambridge Univ. Press, 1990).
- Houghton, J. T. *et al.* (eds) *Climate Change 1995: The Science of Climate Change* (Cambridge Univ. Press, 1996).
- Mitchell, J. F. B., Johns, T. C., Gregory, J. M. & Tett, S. F. B. *Nature* **376**, 501–504 (1995).
- Hasselmann, K. *et al.* *Max-Planck-Institut für Meteorologie Report No. 168* (1995).
- Santer, B. D. *et al.* *Climate Dynamics* **12**, 79–100 (1995).
- Taylor, K. E. & Penner, J. E. *Nature* **369**, 734–739 (1994).
- Ramaswamy, V., Scharzkopf, M. D. & Randel, W. J. *Nature* (submitted).
- Gaffen, D. J. *J. geophys. Res.* **99**, 2667–2676 (1994).
- Parker, D. E. & Cox, D. I. *Int. J. Clim.* **15**, 473–496 (1995).
- Hirst, A. C. & McDougall, T. J. *J. phys. Oceanogr.* **26**, 1320–1343 (1996).

immediately dependent on the duplication and segregation of chromosomes.

The proliferation of tumour cells cannot therefore be driven by mutations that merely deregulate the chromosome cycle, because this would not by itself promote the cellular biosynthesis needed for a tumour cell's increase in mass. The discovery that Rb not only regulates genes involved in driving the chromosome cycle (through E2F), but also rRNA and tRNA genes needed for increasing the cell's biosynthetic capacity, therefore begins to explain how a loss of Rb function might have such a powerful effect on cell proliferation.

We are, alas, still far from fully understanding the role of the Rb pathway in preventing tumour formation. Mouse embryos lacking either p16 or Rb, or over-

expressing cyclin D, develop surprisingly normally^{5,14}. Transcriptional repression implemented by Rb is clearly not essential for controlling most cell proliferation in mammals. Discovering the physiological processes whose alteration (by mutation) allows loss of Rb to become crucial for the growth of human tumour cells is clearly an important goal for the future. Rb is located in the nucleus and has thus far only been implicated in controlling transcription. The increased biosynthesis of tumour cells presumably requires not only mutations to Rb that enhance transcription, but also others that enhance translation of messenger RNA. □

Kim Nasmyth is at the Research Institute of Molecular Pathology, Dr Bohr-Gasse 7, A-1030 Vienna, Austria.

receptor-type tyrosine kinase whose ligand and has remained unknown, but the gene concerned has attracted considerable attention because inherited gain-of-function mutations predispose human carriers to certain cancers (thyroid carcinoma and multiple endocrine neoplasia). Furthermore, loss-of-function mutations in the *ret* gene are found in some familial forms of Hirschsprung's disease, a disorder of neural-crest development that includes defects in intestinal innervation (see refs 3, 4 and 6 for references on Ret).

Given the similarities between the Ret- and GDNF-null phenotypes, various groups³⁻⁶, including some of those reporting on GDNF-defective mice, have asked whether Ret might be a GDNF receptor. The answer in all cases is a resounding "Yes". These papers show that Ret expression confers GDNF-binding activity, that GDNF associates with Ret, and that this association induces Ret phosphorylation on tyrosine, a sign of ligand-induced receptor activation. Moreover, addition of GDNF induces neuroblast axonal outgrowth in nephrogenic explants from normal mice, but not from Ret-defective mice⁴. Consistent with a ligand-receptor relationship, Ret is expressed in epithelial cells in the ureteric bud whereas GDNF is expressed in neighbouring mesenchymal cells⁴. All in all, the case for Ret being a functional GDNF receptor is very strong.

What, then, of GDNFR- α ? Returning to the PIPLC experiments, Treanor *et al.*⁵ and Jing *et al.*⁶ noticed that GDNF failed to bind Ret and induce Ret phosphorylation in cells treated with PIPLC. Both functions could be restored by adding a soluble form of GDNFR- α together with GDNF to the cells treated with PIPLC. In related experiments, overexpression of GDNFR- α augmented GDNF binding to Ret, and the three proteins could be immunoprecipitated in the same complex. In the model suggested by these observations, GDNFR- α binds GDNF, and this complex is subsequently recognized by Ret (see *d* in the figure).

Ligand binding to a signalling receptor through the agency of another molecule is not unusual (*a-c* in the figure). For example, fibroblast growth factor (FGF) binds to tyrosine kinase receptors in association with heparan sulphate glycosaminoglycans, and TGF- β binds to serine/threonine kinase receptors in association with the membrane protein betaglycan¹⁰. A GPI-anchored protein analogous to GDNFR- α mediates binding of ciliary neurotrophic factor (CNTF) to a receptor subunit that activates non-receptor tyrosine kinases¹¹. At the least, the function of these molecules is to limit ligand diffusion from the cell surface¹⁰.

Compared with the dependence of other receptor tyrosine kinases on accessory receptors, however, the dependence of Ret on GDNFR- α is unusually strict.

- Hunter, T. & Plines, J. *Cell* **79**, 573-582 (1994).
- White, R. J., Trouche, D., Martin, K., Jackson, S. P. & Kouzarides, T. *Nature* **382**, 88-90 (1996).
- Hinds, P. W. & Weinberg, R. A. *Curr. Opin. Genet. Dev.* **4**, 135-141 (1994).
- Sherr, C. J. *Trends Cell Biol.* **4**, 15-18 (1994).
- Weinberg, R. A. *Cell* **81**, 323-330 (1995).
- Cavanaugh, A. H. *et al. Nature* **374**, 177-180 (1995).
- Mitchison, J. M. *The Biology of the Cell Cycle* (Cambridge Univ. Press, 1970).
- Nurse, P. *Nature* **256**, 547-551 (1975).
- Nash, R., Tokiwa, G., Anand, S., Erickson, K. & Futcher, A. B. *EMBO J.* **7**, 4335-4346 (1988).
- Kung, A. L., Sherwood, S. W. & Schimke, R. T. *J. Biol. Chem.* **268**, 23072-23080 (1993).
- Killander, D. & Zetterberg, A. *Expl Cell Res.* **40**, 12-20 (1965).
- Ohtsubo, M. & Roberts, J. M. *Science* **259**, 1908-1912 (1993).
- Resnitzky, D., Gossen, M., Bujard, H. & Reed, S. I. *Molec. cell. Biol.* **14**, 1669-1679 (1994).
- Serrano, M. *et al. Cell* **85**, 27-38 (1996).

NEUROTROPHIC FACTORS

Crossing receptor boundaries

Joan Massagué

THE discovery of GDNF (glial-cell-line-derived neurotrophic factor) as a survival factor for dopaminergic neurons in the midbrain¹, made in 1993, raised hopes that GDNF might someday be useful against Parkinson's disease and other neuronal disorders, and touched off an intense search for its receptors as potential therapeutic targets. Because GDNF is related to TGF- β (transforming growth factor- β), one might have predicted that the GDNF receptor would be found among the family of transmembrane serine/threonine kinases that act as receptors for TGF- β and related factors².

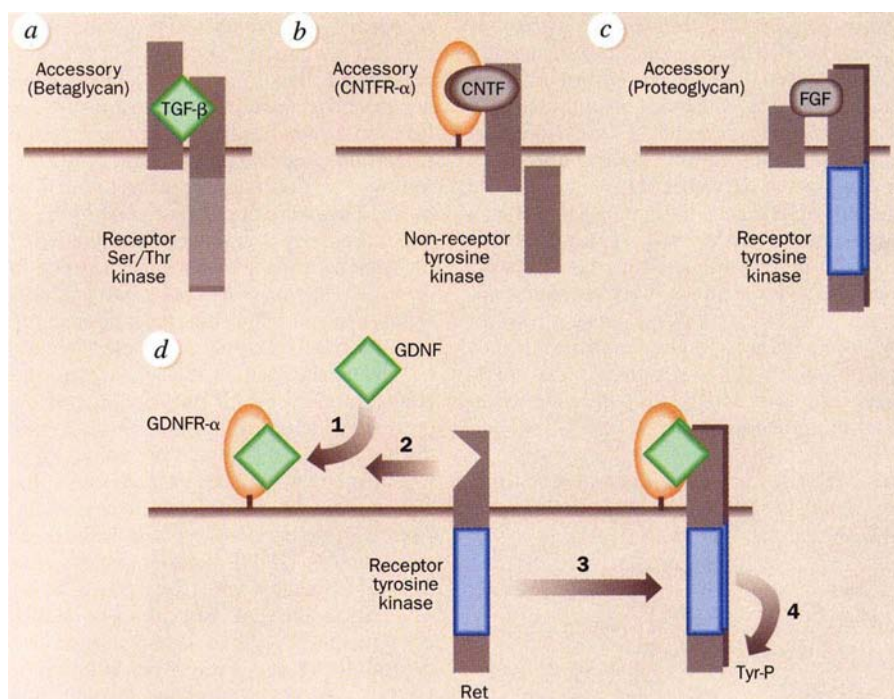
Instead, that search has just turned up something completely different. Three papers in *Nature*³⁻⁵ (two last week^{3,4} and one on page 80 of this issue⁵), and another in *Cell*⁶, identify the receptor tyrosine kinase known as Ret as a functional GDNF receptor. Moreover, two of the groups concerned^{5,6} show that a novel glycolipid-anchored membrane protein, called GDNFR- α , is Ret's partner in ligand binding.

Using expression-cloning strategies to search for GDNF receptors, Treanor *et al.*⁵ and Jing *et al.*⁶ have identified GDNFR- α , a cell-surface protein that binds GDNF and is expressed in GDNF-responsive tissues. GDNFR- α turns out to be an extracellular protein that is tethered to

the plasma membrane by a glycosylphosphatidylinositol (GPI) anchor. Since GDNFR- α lacks transmembrane and cytoplasmic regions, it could not signal on its own. However, when these investigators treated cultured neurons with a phosphatidylinositol phospholipase (PIPLC) that hydrolyses GPI moieties, the cells lost responsiveness to GDNF but not to other neurotrophic factors. This result suggested that GDNFR- α may participate in GDNF signalling by acting as the obligatory partner of a bona fide signalling receptor.

The clue that led to discovery of that signalling partner unexpectedly came from the phenotype of GDNF-deficient mice described in three other papers in this issue⁷⁻⁹. Although GDNF is a potent survival factor for dopaminergic and spinal motor neurons in culture and in lesion models *in vivo*, the development of these neurons in GDNF homozygous null mice appears normal. However, the mice have deficits in certain peripheral neurons and, remarkably, they fail to develop kidneys and to innervate the gastrointestinal tract. This is not entirely surprising because, during mouse development, GDNF is expressed in the kidney and gastrointestinal tract as well as in the nervous system.

As it turns out, a failure to develop kidneys and an enteric nervous system is also typical of Ret-deficient mice. Ret is a



Comparison of signalling complexes of *a*, transforming growth factor- β ; *b*, ciliary neurotrophic factor; *c*, fibroblast growth factor; and *d*, glial-cell-line-derived neurotrophic factor. GDNF is akin to members of the TGF- β family, but generates a complex with components of the other systems: it gains access to the tyrosine kinase Ret with help from a newly discovered entity, GDNFR- α , an extracellular protein that is attached to the plasma membrane by glycosyl phosphatidylinositol. In step 1, GDNF- α binds GDNF and, in 2, the complex is recognized by Ret. Upon formation of a stable complex between GDNF, GDNFR- α and Ret (3), Ret undergoes tyrosine phosphorylation (4), which leads to signalling. The special interest in the search for receptors for GDNF has stemmed from the earlier finding that this factor helps to promote the survival of dopaminergic neurons, failure of which is implicated in Parkinson's disease and other neuronal disorders.

The fact that a soluble form of GDNFR- α still allows GDNF binding to Ret suggests that GDNFR- α may force GDNF into a shape that can be recognized by Ret. The function of GDNFR- α clearly goes beyond limiting diffusion of a ligand: in essence, it appears that the Ret ligand is GDNF bound to GDNFR- α .

What is also remarkable about GDNF is that it is a TGF- β family member, but one that signals through a receptor tyrosine kinase rather than a serine/threonine kinase. This breach of a family tradition may reflect GDNF's status as the most distant TGF- β relative known to date. Its formal inclusion in this family hinges on the facts that GDNF has TGF- β -like precursor structure and dimeric subunit composition, and that it contains a set of cysteines typical of the family¹. These cysteines form three disulphide bonds interlocked into a tight structure known as the 'cystine knot'¹². To be sure, however, the cystine knot is not exclusive to the TGF- β family. Versions of this structure are present in nerve growth factor (NGF) and platelet-derived growth factor (PDGF), both of which signal through typical receptor tyrosine kinases, and in human chorionic gonadotropin, which signals through a G-protein-coupled receptor¹².

The cystine knot is one of the various core structures found in polypeptides that function in the extracellular milieu. As a

group, the receptor tyrosine kinases have evolved to accept a wide range of such structures as ligands. With this in mind, GDNF could be viewed as just another structure recognized by receptor tyrosine kinases (in this case, a TGF- β -like structure). On the other hand, GDNF might be less of a TGF- β family renegade than it seems. Could some TGF- β family members signal through receptor tyrosine kinases as well as through receptor serine/threonine kinases? When the latter receptors took centre stage in this field a few years ago, any thoughts of tyrosine kinases being involved in TGF- β signalling faded away. It may now be time to explore those thoughts again. □

Joan Massagué is in the Cell Biology Program, Howard Hughes Medical Institute, Box 116, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, New York 10021, USA.

1. Lin, L.-F. *et al.* *Science* **260**, 1130–1132 (1993).
2. Massagué, J. *Cell* **85**, 947–950 (1996).
3. Trupp, M. *et al.* *Nature* **381**, 785–789 (1996).
4. Durbec, P. *et al.* *Nature* **381**, 789–793 (1996).
5. Treanor, J. *et al.* *Nature* **382**, 80–83 (1996).
6. Jing, S. *et al.* *Cell* **85**, 1113–1124 (1996).
7. Sánchez, M. P. *et al.* *Nature* **382**, 70–73 (1996).
8. Pichel, J. G. *et al.* *Nature* **382**, 73–76 (1996).
9. Moore, M. W. *et al.* *Nature* **382**, 76–79 (1996).
10. Schlessinger, J. *et al.* *Cell* **83**, 357–360 (1995).
11. Stahl, N. & Yancopoulos, G. D. *Cell* **74**, 587–590 (1996).
12. Sun, P. D. & Davies, D. A. *Rev. Biophys. biomolec. Struct.* **24**, 269–291 (1995).

DAEDALUS

Free-flowing salt

THE electrochemical salt bridge is an open glass tube containing a solid agar gel loaded with salt. The ions in the gel are quite mobile, so the bridge allows current to flow between two solutions while not allowing them to mix.

Daedalus is now generalizing this idea. Some ion-exchange resins are said to conduct electricity feebly; their absorbed ions can hop through the resin from one charged site to another. He suspects that in fact they are excellent conductors. Their high resistance arises only at the metal terminals of the test meter; the ions cannot travel through metal or transfer their current to it.

So DREADCO physicists are extruding ion-exchange resins into wire, and making transformer windings from it. Joining the ends will give a complete resin circuit, with no metal anywhere to block the current. As a secondary winding to a conventional transformer, the resin should be able to carry huge currents, maybe even approaching superconducting intensities — especially at high frequencies. For the higher the frequency, the shorter the distance each ion has to run before reversing its path. Indeed, at frequencies of many megahertz, the heavy individual ions will merely be shifting their weight from one foot to the other, so to speak. Scattering, the major source of resistive loss, will be almost negligible.

At the same time, DREADCO chemists are seeking better resins. Their models are the zeolite molecular sieves, and compounds such as urea and polyvinyl alcohol, with long canals through their structure capable of accepting guest molecules and ions. Their goal is a highly crystalline parallel-chain polymer with channels down which ions can flow as freely as possible. Cunningly, they are choosing long, thin ions (such as azide and triiodide), slightly mismatched to the lattice spacing of the polymer. They will find no well-fitting sites on which to bind, so they will move freely and losslessly through the lattice.

The final product should be a set of resin conductors as good as metal. Novel semiconductors, inhabited not by mere holes and electrons, but by a whole rich lattice-fauna of ions, should be possible. An ionic channel in a resin lattice could be blocked by one molecule sitting on the opening, thus creating an ultra-sensitive new chemical or biochemical sensor. Indeed, ionic resins are closely analogous to nerve membranes; and Daedalus wonders what the new ionic currents will feel like. Instead of the usual unpleasant electric shock, they may induce far more informative and rewarding sensations.

David Jones

NO sexual behaviour in newts

SIR — The interaction of nitric oxide synthase with L-arginine generates nitric oxide and L-citrulline¹. Nitric oxide exhibits many regulatory effects in many tissues and systems²⁻⁴, including emotion-regulating brain regions⁵⁻⁷. Nitric oxide regulates sexual behaviour in mammals^{8,9}, but nothing is known about its effects in lower vertebrates. Here we show that nitric oxide has a role in male courtship behaviour but not in females of the urodele crested newt (*Triturus cristatus*).

This species is a good model because it shows a complex, peculiar courtship¹⁰⁻¹². The male is the active partner during courtship, which is in four phases: (1) approach, in which the male sniffs the female's head; (2) fan, in which the male slowly beats its tail towards the female's head; (3) lashes, in which the male's tail hits the female's head; and (4) deposition, in which the male shows his tail tip, the female touches his tail, he deposits the spermatophore, and she picks the spermatophore up with her cloaca. Males do not always complete courtship; in the

case of females, receptive individuals remain immobile and close to the partner until spermatophore deposition, whereas non-receptive ones move away from the courting male.

We captured newts during the reproductive period, put one male and one female together in aquaria and observed the males' courtship. We divided the males into experimental groups: inactive (males which in the presence of a female did not exhibit courtship for at least 6 h), approach, fan, lashes, deposition, post-deposition (15 min after male courtship interruption), and female interruption (0 and 15 min after female moving away). We determined nitric oxide (NO) synthase (NOS) activity *in vitro* in the brain by the methods described in the figure legend.

Inactive males had the lowest values of brain NOS. The activity of this enzyme progressively and significantly increased during the phases of courtship, reaching a maximum during spermatophore deposi-

tion. Males killed 15 min after deposition had brain NOS levels as low as those of inactive males. Brain NOS values in males that interrupted courtship were significantly lower than those of courting males. Males whose partner had just moved away had brain NOS values similar to those of males with receptive females, but 15 min after interruption, these values had decreased to those of inactive newts. Female brain NOS activity did not change in any experimental group.

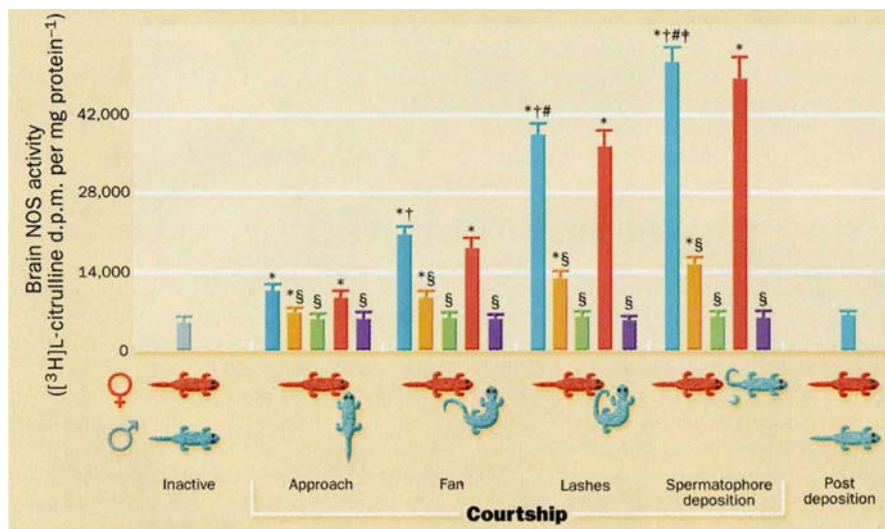
Thus there is a causal relationship between male sexual behaviour and brain NO, in agreement with observations in male mice, which exhibit inappropriate sexual behaviour and increased aggressiveness on disruption of neuronal NOS⁹. In our male newts, reproductive success is characterized by high brain NOS activity which progressively increases over the courtship various phases, just the opposite of inactive newts, which have the lowest brain NOS activity. The passage from one phase to the next seems to require a brain NO 'threshold' value to be reached. This suggestion is supported by observations that the males which interrupt courtship show a NOS activity significantly lower with respect both to those which completed the courtship phase and to those whose female partner moved away.

Female 'receptivity' and 'non-receptivity' do not depend on brain NO; in fact, NOS did not change either when the female remained close to the partner or when she moved away. Unlike males, these results in newts are not in agreement with data from female rats⁸. Many external and internal stimuli influence courtship behaviour. In *T. cristatus*, NO is probably an internal cerebral stimulus, suggesting that its involvement in reproductive behaviour could be conserved throughout vertebrate evolution.

Massimo Zerani

Anna Gobetti

Department of Molecular, Cellular and Animal Biology,
University of Camerino,
via F. Camerini 1, 62032 Camerino MC,
Italy



Brain NOS activity values of male *T. cristatus* during sexual behaviour. Five males in each experimental group were killed by decapitation. NOS activity was determined in the brains by monitoring the conversion of [³H]L-arginine into [³H]L-citrulline, with a modified method previously described¹³. Each brain was weighed, homogenized in 1 ml cold fresh homogenizing buffer (50 mM Tris, 1 mM EDTA, 1 mM EGTA, pH 7.4), and centrifuged at 20,000g for 60 min at 4 °C. Supernatant (25 µl) and incubation buffer (100 µl; 1.5 mM NADPH, 1 mM CaCl₂) containing 150,000 d.p.m. [³H]L-arginine were added to the incubation tube. After 30 min, the enzymatic reaction was stopped by addition of 2 ml blocking buffer (20 mM HEPES, 2 mM EDTA, pH 5.5). The mixture was applied to a pre-equilibrated column (20 mM sodium acetate, 2 mM EDTA, 0.2 mM EGTA, pH 5.5) containing 1 ml Dowex AG50W-X8, and the material eluted with 2 ml water. [³H]L-Citrulline was quantified in a liquid scintillation system. Additional determinations were performed in the presence of excess NOS inhibitor (L-NAME) to verify the specificity of the assay for production of [³H]L-citrulline by NOS. Each mean refers to five determinations ± s.d. One-way analysis of variance $F_{(21,88)}: 12.34, P < 0.01$. pale blue box, inactive; blue box, complete courtship; orange, immediately after courtship interrupted by male; green box, 15 min after courtship interrupted by male; red box, immediately after courtship interrupted by female; purple box, 15 min after courtship interrupted by female. Symbols: Duncan's multiple range test: * $P < 0.01$ versus 'inactive' and 'post-deposition'; † $P < 0.01$ versus 'approach'; # $P < 0.01$ versus 'fan'; § $P < 0.01$ versus 'lashes'; § $P < 0.01$ versus same phase of complete courtship.

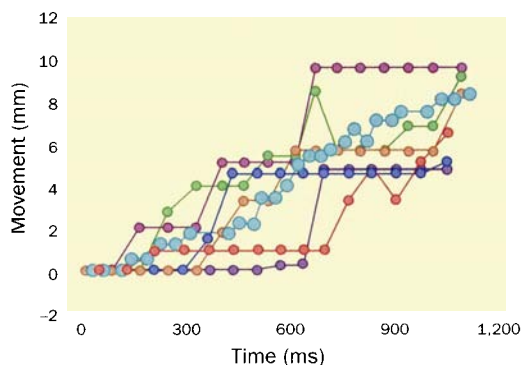
- Nathan, C. & Xie, Q.-w. *Cell* **78**, 915-918 (1994).
- Xie, Q.-w. et al. *Science* **256**, 225-228 (1992).
- McCall, T. & Hassid, A. *Trends pharmac. Sci.* **13**, 1-6 (1992).
- Bredt, D. S. & Snyder, S. H. *Neuron* **8**, 3-11 (1992).
- Bredt, D. S. et al. *Neuron* **7**, 615-624 (1991).
- Vincent, S. R. & Kimura, H. *Neuroscience* **4**, 755-784 (1992).
- Rodrigo, J. et al. *Phil. Trans. R. Soc.* **B345**, 175-221 (1994).
- Mani, S. K. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 6468-6472 (1994).
- Nelson, R. J. et al. *Nature* **378**, 383-386 (1995).
- Halliday, T. R. in *The Reproductive Biology of Amphibians* (eds Douglas, D. H. & Guttman, S. I.) 185-232 (Plenum, New York, 1977).
- Zerani, M. et al. *Horm. Behav.* **26**, 56-61 (1992).
- Andreoletti, G. E. et al. *Horm. Behav.* **17**, 103-110 (1983).
- Bush, P. A. et al. *Biochem. biophys. Res. Commun.* **185**, 960-966 (1992).

One-strided waggle dance in bees

SIR — Honeybees recruit nestmates to food sources by waggle dances, which consist of the waggle run and the return run¹. High-speed video recordings now show that the continual forward motion of the bee during the waggle run masks the true nature of this phase of the dance in that the entire 'run' consists of only a single slow stride. This stride is performed in a strange gait in which all six legs hold onto the comb most of the time, regardless of

trast to the waggle 'runs', show clear, coordinated patterns in the sequence of leg movements (tripod gait). Another difference lies in the ratio of the lift-off to the ground-contact phases. This ratio for the waggle 'runs' is about 1:5, which is characteristic of very slow walks in insects². For a regular walk across the comb the ratio is about 1:2.

Most remarkable, however, was our finding (without exception in 80 dances by 20



Forward movement of the bee's body (head) and the legs during single waggle 'runs' for a feeding station 1,200 m away from the hive. The head moves continuously forward (cyan). The individual tarsi make 1–3 rapid steps but hold fast to the comb for the greater part of the waggle run. The entire waggle run is achieved in a single stride in that at least one of the six legs has moved forward only once. Key: purple, front leg, left; green, middle leg, left; mauve, hind leg, left; cyan, head; blue, front leg, right; brown, middle leg, right; red, hind leg, right.

the distance to the food source.

We recorded dances at 200 frames per second (NAC HSV 400) performed on the surface of the combs in observation hives by marked foragers of honeybees (*Apis mellifera carnica* Pollm.). The foragers were trained to two feeding stations, 200 and 1,200 m from the hive. These distances were chosen because they represent two extremes in the relationship between the distance to a feeding site and the resulting length and pace of the waggle run¹.

We analysed the position of the head and tarsi of the dancing bee frame by frame. The dancing bee moves continuously and slowly forward, covering a distance of about 5 mm for the feeding station 200 m away, and about 8 mm for the station 1,200 m away. The speed of the forward movement is different for the two feeding stations, being faster for the nearer one. The movements of the legs during the forward motion of the body in the waggle are very different from the normal tripod pattern expected of insect walking. During the waggle 'run', individual legs move in small (sometimes uncertain) steps whereas others remained fixed to the edges of the honeycomb cells. There is no orderly pattern of temporal coordination between the legs. Most of the time, all six, five or four legs remained fixed to the comb during the waggle (see figure).

To exclude the possibility that this curious gait may be due to the difficulty of moving across a surface consisting essentially of an open-meshed grid, we analysed the gait of non-dancing bees moving across open-celled comb and across a smooth surface. Both such walks, in con-

bees for both distances) that the entire waggle 'run', regardless of its length, consists of a single stride. For the foragers returning from a source 1,200 m away, this single stride takes 1.2–1.8 s ($n=40$) to cover 8 mm. A single stride over a smooth surface also covers about 8 mm, but takes only 80–120 ms ($n=10$). The simplest interpretation of

our finding is that the bees are trying to maximize their stability on the comb during the waggle movement and thus a transmission of vibrations into the comb during their dance. Why should they do this?

Honeybees nesting in the open can perform long-lasting waggle dances over distances of several centimetres³. These dances are performed not on the comb but on the bodies of other bees³, which means that substrate vibration cannot be a communication channel. *Apis mellifera*, on the other hand, communicates in the darkness of an enclosed nest. Here, comb vibrations seem to be an important communication factor⁴. The waggle movement with feet kept in place may simply be a mechanical 'trick' to increase the transfer of pulsed vibrations that occur during the waggle run⁵ as a communication signal into the comb.

Jürgen Tautz

Kristin Rohrseitz

Theodor Boveri Institut,

Universität Würzburg,

Am Hubland,

D-97074 Würzburg, Germany

David C. Sandeman

School of Biological Science,

The University of New South Wales,

PO Box 1, Kensington,

New South Wales, Australia 2033

1. von Frisch, K. *Tanzsprache und Orientierung der Bienen* (Springer, Berlin, 1965).
2. Manton, S. M. *Habits, Functional Morphology and Evolution* (Oxford Univ. Press, 1977).
3. Dyer, F. C. in *Neurobiology and Behavior of Honeybees* (eds Menzel, R. & Mercer, A.) 54–65 (Springer, Berlin, 1987).
4. Tautz, J. *J. exp. Biol.* **199**, 1375–1381 (1996).
5. Esch, H. Z. *vergl. Physiol.* **45**, 1–11 (1961).

Prion phylogeny revisited

SIR — Krakauer *et al.*¹ presented a phylogeny of prion (PrP) protein and concluded that the presence of two amino-acid replacements uniquely shared by hominoids (H) and cattle (C) [asparagine-to-serine in codon 143(H), 146(C) and tyrosine-to-histidine in codon 155(H), 158(C)] has biological significance, as such a situation seems unlikely to have occurred by chance. They also speculated that the shared changes may have predisposed humans towards susceptibility to BSE (bovine spongiform encephalopathy). Their conclusions, however, ignore relevant experimental evidence.

The argument of Krakauer *et al.*¹ is based on the hypothesis that similarity between PrP protein sequences is linked to susceptibility because of enhanced interaction between donor and recipient PrP proteins. However, experimental transmission of BSE to mice² and sheep³ suggests that susceptibility is not controlled by overall sequence similarity between donor and recipient PrPs, as sheep encoding a PrP protein variant differing from bovine PrP in codons 143[H]

(146[C]), 155[H] (158[C]) and four other positions (98[C], 100[C], 189[C], 208[C]), succumb to experimental BSE challenge in approximately 500 days³.

This level of susceptibility is similar to that seen following challenge of sheep with scrapie, where there is no PrP sequence difference. However, sheep that are heterozygous for a PrP gene encoding the additional substitution of arginine for glutamine at codon 171[C] do not develop BSE or scrapie until at least 1,800 days after challenge, and sheep homozygous for arginine 171 appear resistant³. Thus, the combined effects of six amino-acid differences between sheep and bovine PrP proteins are apparently outweighed by a single additional change in a crucial position at codon 171. Only humans (apart from sheep) have glutamine replaced at this position 168(H), this time with glutamate. Because the cattle PrP gene has glutamine at this position, perhaps transmission of BSE to humans is less likely as a consequence. However, even this suggestion is highly speculative.

There are species to which BSE appears to have been transmitted via con-

taminated food⁴ and which do not share the codon 146(C) and 158(C) amino acids with cattle: oryx⁵ and domestic cat (W. G., unpublished data). Many amino-acid changes in PrP change the phenotype of prion diseases; it is intriguing to speculate about evolutionary adaptation against these diseases operating via sequence changes in PrP protein. It may therefore be the case that codons 143, 155 and, as suggested above, 168 of the human PrP gene are key amino-acid positions in the interaction with components of the infective agent. However, a correlation between a particular feature of a PrP protein sequence and predisposition to a certain strain of agent can only be demonstrated experimentally.

In addition, Krakauer *et al.* based their analysis on only 56 variant codon positions. In fact, there are at least 80 positions in the full open reading frame which show a minimum of one amino-acid replacement, leading to 3,160 possible different pairs of amino-acid replacements. There are, for example, eight sheep PrP protein variants and not the single sheep sequence shown in the phylogeny⁶. We suggest that a calculation on the basis of all prion protein sequences (including all known polymorphisms) would not only change the phylogenetic tree but would result in a higher probability for the occurrence by chance of the convergence observed. Serine and asparagine are highly variable in many codon positions of PrP. In addition to hominoids and cattle, the codon 143 asparagine-to-serine replacement occurs twice as a single replacement: in pig⁷, omitted by Krakauer *et al.*¹, PrP analysis also shows that codon 155 is one of the few positions which have two possible amino-acid replacements (asparagine in hamster), perhaps indicating a tolerance for differences in this region.

A priori prediction of the incidence or cross-species transmissibility of a spongiform encephalopathy requires a more adequate three-dimensional structural model for PrP protein and its molecular interactions than is available, and a much better understanding of the molecular basis of the strains of these unconventional disease agents.

Wilfred Goldmann

Nora Hunter

Robert Somerville

James Hope

BBSRC & MRC Neuropathogenesis Unit,
Institute for Animal Health,
West Mains Road, Edinburgh EH9 3JF, UK

1. Krakauer, D. C., Pagel, M., Southwood, T. R. E. & Zanotto, P. M. de A. *Nature* **380**, 675 (1996).
2. Bruce, M. *et al.* *Phil. Trans. R. Soc.* **B343**, 405–411 (1994).
3. Goldmann, W., Hunter, N., Smith, G., Foster, J. & Hope, J. J. *gen. Virol.* **75**, 989–995 (1994).
4. Wilesmith, J. W., N. Z. *Vet. J.* **42**, 1–8 (1994).
5. Poudinger, M., Kirkwood, J. & Almond, W. *Archs Virol.* **131**, 193–199 (1993).
6. Hunter, N. *et al.* *Archs Virol.* (in the press).
7. Martin, T., Hughes, S., Hughes, K. & Dawson, M. *Biochem. biophys. Acta* **1270**, 211–214 (1995).

Vertebrate with protrusible eyes

SIR — Caecilians are among the most divergent and least-known major vertebrate groups. Found throughout the humid tropics (with the exception of Madagascar), these limbless, fossorial

African family Scolecomorphidae, the eye is not under bone or in a socket. Instead, it is attached laterally to the base of the tentacle^{7–9}. In most caecilians, this would prevent the eye from being exposed to light,



a, Living specimen of *S. kirkii* with the tentacle retracted. Note that the eye is still visible in the tentacular groove under a layer of unpigmented skin. *b*, Another specimen with the tentacle completely protracted, clearly showing the eye beyond the roofing bones of the skull.

amphibians possess many unique and often strange characteristics, not the least being their tentacles. These protrusible organs, one on each side of the snout, are derived from the tear duct, certain extrinsic eye muscles and other parts of the eye^{1,2}. The tentacle can be protracted a considerable distance out of the head in many species. Anatomically, it is directly connected to the vomeronasal organ³, and behavioural experiments suggest that it is involved in chemoreception⁴. We now show for the first time that the east African caecilian *Scolecophorus kirkii* can protrude its eyes beyond the skull by means of the tentacle.

In all caecilians, the eyes are lidless and reduced, and at best function to detect changes in light intensity^{5,6}. In many species the eyes are fixed in bony sockets, in others they are completely covered by the roofing bones of the skull, while yet others apparently lack eyes entirely. In most caecilians that have eyes, the eye and tentacle are morphologically distinct, the tentacle residing in a separate opening in the skull called the tentacular groove. However, among the species of the

as it would be under heavily pigmented skin. But in scolecomorphids, the tentacular groove is covered by pigmentless skin. Thus, as the tentacle is protracted and retracted, the eye moves with it along a translucent track and remains exposed to ambient light regardless of position (*a* in the figure). The anatomy of preserved museum specimens of scolecomorphids suggested that their eyes might be protrusible⁹. That is, these animals might be able to protract the tentacle far enough to move the eye beyond the roofing bones of the skull, out through the tentacular aperture and into the external environment. Several living specimens of *S. kirkii* became available recently, giving us the opportunity to test this hypothesis. Using a 35-mm camera and a synchronized strobe, we captured images (*b* in the figure) which demonstrate that *Scolecophorus* can indeed protrude its eyes beyond the skull. Some ver-

tebrates (for example, mudskippers and frogs) have eyes that are retractable within their sockets, but this is the only known vertebrate with highly mobile, protrusible eyes.

James C. O'Reilly

Department of Biological Sciences,
Northern Arizona University,
Flagstaff, Arizona 86011-5640, USA

Ronald A. Nussbaum

Museum of Zoology,
University of Michigan,
Ann Arbor, Michigan 48109-1079, USA

Daniel Boone

Bilby Research Center,
Northern Arizona University,
Flagstaff, Arizona 86011-6013, USA

1. Wiedersheim, R. *Die Anatomie der Gymnophiona*, 44–57 (Fischer, Jena, 1879).
2. Billo, R. & Wake, M. H. *J. Morph.* **192**, 101–111 (1987).
3. Badenhorst, A. *Ann. Univ. Stellenbosch (A2)* **1**, 1–26 (1978).
4. Himstedt, W. & Simon, D. *Herpet. J.* **5**, 266–270 (1995).
5. Wake, M. H. *Zoomorphology* **105**, 277–295 (1985).
6. Fritsch, B., Himstedt, W. & Capron de Caprona, M.-D. *Dev. Brain Res.* **23**, 201–210 (1985).
7. Boulenger, G. A. *A. Mag. nat. Hist.* **11**, 48 (1883).
8. Nussbaum, R. A. *Copeia* **1981**, 265–269 (1981).
9. Nussbaum, R. A. *Occip. Pap. Mus. Zool. Univ. Michigan* **713**, 1–49 (1985).

Environmental Trifluoroacetate

SIR — Partially fluorinated ethanes (HFCs) are being introduced as coolants for refrigeration and air conditioning as alternatives for the ozone-depleting chlorofluorocarbons. For 1,1,1,2-tetrafluoroethane, the most important HFC (134a), an annual production of 230,000 metric tons is expected for the year 2010¹, 30–40 % of which will be atmospherically oxidized to trifluoroacetic acid (TFA)^{2,3}. Based on the assumption that there is at present no environmental TFA, levels of 5–20 pg m⁻³ in air and of 100–160 ng l⁻¹ in precipitation have been forecast for 2010^{4,5}. However, air and water samples collected in Germany, Switzerland and Israel in 1995 already contain levels of TFA in the range predicted for 2010.

We collected air samples (30–60 m³) using denuder tubes with alkaline coating at a rate of 0.6 m³ h⁻¹. We determined TFA by GC-MS⁶, always in triplicate. The air concentrations in Bayreuth range from

20 to 150 pg m⁻³ (see table), at Waldstein from 20 to 40 pg m⁻³, and in Zurich are about 65 pg m⁻³ (data not shown).

Similarly, we found that TFA levels in rain water are 25–280 ng l⁻¹ in Bayreuth (see table), and in Zurich and at Alpthal (Switzerland) are 40–80 ng l⁻¹ (data not shown). The Roter Main River (Bayreuth) contained 60–280 ng l⁻¹ TFA, similar to the levels in other rivers and lakes in Germany. TFA in spring waters in the Fichtelgebirge mountains in Germany are 70–320 ng l⁻¹. Lake Kinneret (Israel) and the Jordan River have concentrations of 200 and 250 ng l⁻¹, respectively, and the Dead Sea 6,400 ng l⁻¹. The last figure is perhaps an example of evaporative accumulation, as discussed previously⁷.

Similar concentrations of TFA in precipitation and in spring water indicate that it is neither degraded nor absorbed to the soil and may reach the ground water, although it has been suggested that TFA

may be metabolized anaerobically in soils⁸. The ages of the respective spring waters are unknown; therefore, further studies are required to determine the origin of TFA in these samples.

The total amount of TFA in precipitation on Germany is about 30 metric tons annually (average precipitation 800 mm, surface area 357,000 km²). The annual flow of TFA from the four main rivers of Germany to the sea (flow rates: Rhine, 2,170 m³ s⁻¹; Weser, 360 m³ s⁻¹; Elbe, 815 m³ s⁻¹; Danube 1,420 m³ s⁻¹) is about 55 tons per year. Because half of the Rhine and Danube water originate in other countries, TFA export by rivers from Germany corresponds to wet deposition.

The inhalation anaesthetic halothane (1,1,1-trifluoro-2-bromo-2-chloro-ethane) is a major precursor of TFA; about 870 metric tons a year are deposited throughout the world⁹. The atmospheric lifetime of halothane is about a year, so TFA formed by this route should be evenly distributed globally; less than 0.1 % would be expected to be deposited on Germany. The estimated global TFA rainwater level from this source and from tropospheric HFC-134a¹⁰ seems to be about 3 ng l⁻¹. Thus, relatively high average concentrations of TFA in rain (100 ng l⁻¹) suggest that there are additional, unknown sources.

Ocean surface waters have TFA concentrations of 40–250 ng l⁻¹ (see table). Although this indicates a potentially large environmental inventory of TFA, more data on oceanic depth profiles are needed before we can reach any conclusions about the origin of this TFA.

Hartmut Frank

Armin Klein

Dirk Renschen

*Department of Environmental Chemistry
and Ecotoxicology,
University of Bayreuth,
D-95440 Bayreuth,
Germany*

TFA-CONCENTRATIONS IN AIR, PRECIPITATION AND SURFACE WATERS IN 1995						
	Air (pg m ⁻³)	Rain (ng l ⁻¹)	Surface waters (ng l ⁻¹)			
			Roter Main	Rivers / Springs	Lake/Seawater	
March	55 ± 20 (n = 3)	100 ± 50 (n = 8)	160 ± 27 (n = 3)	215 MB 200 D 150 P		
April	70 ± 40 (n = 9)	40 ± 4 (n = 4)	80	260 N		
May	50 ± 15 (n = 5)	130 ± 65 (n = 6)	140	130 S Sp 110 E Sp 80 Nb Sp 70 WM Sp 320 F Sp	70 LF 115 LW	
June	50	190; 240	110			
July	25; 25	70 ± 20 (n = 5)		100 E 100 W 630 R 60 Ww	90 NS 40 BS	
Aug.		120 ± 110 (n = 3)				
Sept.	70; 120	100			60 LC 250 Atl.F	
Oct.	30	75	280		70 Atl.I	
Nov.	20	60 ± 55 (n = 3)	110	250 J	200 LK 6,400 DS	
Dec.	20	30	60			
Average	50	100	140			

All values are corrected for blank contribution. Three blank samples were always prepared in parallel; blank TFA concentrations were 10 ± 3 ng/L, giving a limit of determination of 20–40 ng l⁻¹. For air the limit of determination is 20 pg m⁻³ (air sample volume = 50 m³). Sampling sites: air/precipitation/Roter Main: Bayreuth (360 m a.s.l.; E:11°35', N:49°56'); Surface waters: MB, Mistelbach (Bayreuth); D, Danube (Regensburg); P, Pegnitz (Nürnberg) N, Neckar (Tübingen); S Sp, Saale Spring (Fichtelgebirge mountains, FG); E Sp, Eger Spring (FG); Nb Sp, Naab Spring (FG); WM Sp, White Main Spring (FG); F Sp, Friesen Spring (FG); E, Elbe (Hamburg-Altona); W, Weser (Bremerhaven); R, Rhine (Duisburg-Ruhrort); Ww, Warnow (Rostock); J, Jordan (shortly below Lake Kinneret/Israel); LF, Lake Fichtel (FG); LW, Lake Weissenstadt (FG); NS, North Sea (Husum); BS, Baltic Sea (Warnemünde); LC, Lake Constance (Constance); Atl.F, Atlantic (Ile d'Yeu/France); Atl.I, Atlantic (Mace Head/Ireland); LK, Lake Kinneret (Capernaum/Israel); DS, Dead Sea (En Bokek/Israel).

- McCulloch, A. *Env. Int.* **21**, 353–362 (1995).
- Wallington T. J., Hurley, M. D., Ball, J. C. & Kaiser, E. W. *Environ. Sci. Technol.* **26**, 1318–1324 (1994).
- Rodríguez, J. M., Ko, M. K. W., Sze, N. D. & Heisey, C. W. *Proc. AFEAS Workshop on Atmospheric Wet and Dry Deposition of Carbonyl and Haloacetyl Halides* 25–32 (AFEAS, Brussels 1992).
- Franklin, J. *Chemosphere* **27**, 1565–1601 (1993).
- Kotamarti, V. R. et al. (Abstr.) *Eos* **75**, 137 (1994).
- Frank, H., Renschen, D., Klein, A. & Scholl, H. J. *High Res. Chromatogr.* **18**, 83–88 (1995).
- Tromp, T. K., Ko, M. K. W., Rodríguez, J. M. & Sze, N. D. *Nature* **376**, 327–330 (1995).
- Visscher, P. T., Culbertson, C. W. & Oremland, R. S. *Nature* **369**, 729–731 (1994).
- McCulloch, A. *AFEAS Workshop on Environmental Fate of Trifluoroacetic Acid* (ed. Chumley, T. G.) Ch. 4 (AFEAS, Miami Beach, 1994).
- Montzka, S. A. et al. *Geophys. Res. Zett.* **23**, 169–172 (1996).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. Priority will be given to letters of fewer than 500 words and five references.

Cases of conviction

Walter Gratzer

Confronting the Experts. Edited by Brian Martin. State University of New York Press, Albany, New York: 1996. Pp. 204. \$14.95 (pbk).

IN a struggle between you and the world, said Kafka, back the world. Those who struggle and pay the price probably earn respect and admiration, more often than obloquy. Yet all too commonly such characters are animated by the simple desire to get a knee in the crotch of their betters, and this does not in itself necessarily lead to the greater good. The study of how to win friends and cajole sceptics is alien to their natures; the soft answer that turneth away wrath is not an option for them. Consider for instance Dharendra Sharma, an opponent of nuclear power policy in India, who plumes himself thus in Brian Martin's collection of zealots: "Temperamentally, I am at my best in confrontations with powerful adversaries". I felt at this the faintest flicker of sympathy for the powerful adversaries.

Martin's contributors are a varied lot; most must be uncomfortable colleagues and one or two are men at whose approach cripples might cast off their crutches and head for the hills. Some appear wholly admirable and have sought only to fight what they saw as dangerous error, for the public good. Notable among these is Sharon Beder, an engineer by training, who chose for her PhD project to examine how the Water Board of her native Sydney had come to opt for a sewage disposal system based on deep-water outfalls. The policy of the Water Board and of the State Pollution Control Commission, it emerged, had been decided by engineers, a calling that, if Beder is to be believed, extols solidarity and views open criticism as culpable disloyalty: engineers, she says, do not disparage other engineers. So the evidence that the sewage would not always sink and disperse, but would too often form floating islands and drift back onto the beaches was suppressed. Beder gained access to forbidden documents and fought her way through a morass of evasion and obfuscation by politicians and others.

The board's consultants, Beder found, purported to have looked into health hazards, although the grounds for proclaiming the beaches to be safe turned out to be the absence of coliform bacteria. But coliforms die off quickly in salt water; they were studied, the report explained, "as a matter of convenience". Viruses, the real

threat, were not so conveniently studied and were accordingly disregarded. Beder persisted and illuminated a dark area of public concern, that seems now finally to have been opened to rational debate.

Of the other contributors, Edward Herman, a professor (now emeritus) at the University of Pennsylvania, took on the biggest adversary of all, the US government. His theme is terrorism and his thesis that, as implicitly defined by the pseudo-

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Miracle machinery? A fake perpetual motion machine invented by *Nature's* own charlatan, David Jones (aka Daedalus), and built by DREADCO wizardry. On display at the Heureka Science Centre, Helsinki, it has been running non-stop for five years.

experts — mostly journalists and those commonly associated at one time or another with organizations such as the Central Intelligence Agency (CIA) — it is misleadingly represented. From 1968 to 1980 the number of assassinations worldwide by all terrorist bodies listed as such by the CIA was 3,680. During 1978–83 alone, twenty times that number of Guatemalan citizens fell victim to state-directed terrorism, with the connivance and even practical help of the US government. Herman writes with authority and eloquence, but is it really true that views like his do not get a hearing? His writings on the subject have been censored, he tells us, well-disposed editors have been sacked and television

appearances cancelled. Yet Noam Chomsky, with whom he has jointly published books and articles, surely receives wide attention, if not adulation. His views may not please the US government, but if it has sought to suppress them, its efforts have been in vain.

So perhaps Herman is none too skilful at persuasion, and the same may well be true of some of the more sanguine of Martin's other authors. None believe that they have received proper credit for their unsleeping endeavours. Mark Diesendorf, who campaigned against fluoridation of drinking water, and Dharendra Sharma see themselves as solitary warriors, lions in a den of Daniels. They are as sure as can be that they are right, and they give the impression of being uncommonly pleased with themselves. They may indeed be right, and if so I salute them; but I have not seen the case for the defence, nor would I feel qualified, without much deeper study, to form a judgement. The same holds true for the art historians Michael Malory and Gordon Moran, who tell a compelling tale. They found archival and structural evidence — their case, as presented, again seems unanswerable and has aroused fearsome passions in Italy — that the equestrian figure of the *condottiere* Guido Riccio in Simone Martini's famous Siena fresco is a later edition. But, as to the biochemist in this company, Harold Hillman, I do feel that I can make a judgement, which is that he is deluded on all counts.

Hillman has a messianic conviction that more or less everyone in the profession is wrong about pretty much everything that matters. He believes for instance that all attempts at cellular fractionation generate nothing but artefacts, that practically everything revealed by the electron microscope is an artefact of fixation and staining, that there is no such thing as a lipid bilayer membrane, that the known cytoplasmic organelles are mirages, that nervous tissue consists of granular "ground substance" with "naked nuclei", and for good measure that ATP is hydrolysed by visible light (which of course it does not absorb). Hillman asserts that the Establishment has conspired to suppress these uncongenial facts, for the mandarins fear the truths that he has exposed — truths that would discredit their own work.

It is Martin's espousal of Hillman's cause that makes me uneasy about some of the rest. For academic freedom is all very well, but it has its price. A couple of decades ago, a number of like-minded scholars launched a journal of science dedicated to free publication without censorship: papers could be on any topic; the

originality and intellectual daring of which no hidebound referees would be permitted to curb. But soon the journal became a magnet for monomaniacs and cranks who had failed to get into print elsewhere. Accounts of perpetual motion machines and extraterrestrial beings flooded into the editorial office. The editor soon realised that some compromise was inescapable: he would filter out the worst excesses, and if he was unsure he would consult a friend or two. I believe the journal, for all its noble ideals, soon vanished.

Everyone in science and no doubt in all other academic disciplines learns to avoid the Ancient Mariners of their profession. They are noisy, manically persistent and they consume time and tax tempers. In mediaeval times they would probably have been dealt with by an instrument called the Holy Maul, which was kept behind the church door and used to brain the aged and feeble-minded when their antics became too tiresome to bear.

Martin's introduction and summing-up are thoughtful and undogmatic, but contain what seem to me some dubious propositions. He believes in the overwhelming power of the Establishment to persuade, but is this really so irresistible? A journalist recently observed that the question that he and his colleagues should ever keep in their minds when interviewing a representative of officialdom is: "Why is this lying bastard lying to me?" The public probably disbelieves most official pronouncements, especially those presented with the greatest sincerity and fervour. There is no fury, it has been said, like a vested interest masquerading as a moral principle. So when Matthew Meselson made fools of the US Navy over the supposed safety of disposing of radioactive waste in sunken ships, or Richard Feynman pulled off his coup at the Challenger disaster inquiry, we all cheered. As to science, though, one might ask how an Establishment is to be recognized. According to Martin, "Hillman published many papers critical of biological orthodoxy, but for many years it appeared that no one took any notice. Only an establishment can get away with this." He does not understand that this supposed Establishment is simply you and me. Most papers and grant applications are reviewed and journals edited by foot-soldiers, not by generals. Outsiders and eccentrics, moreover, *do* get a hearing; and eventually, when some new and revolutionary notion has been tempered in the furnace of scepticism and doubt, it does supplant the old orthodoxy. More often, however, a seeming crackpot is not Einstein or Darwin, but merely a crackpot. □

Walter Gratzer is in the MRC Muscle and Cell Motility Unit, King's College, University of London, 26-29 Drury Lane, London WC2B 5RL, UK.

Farming for beginners

Bernard Wood

The Origins and Spread of Agriculture and Pastoralism in Eurasia. Edited by David Harris. UCL Press: 1996. Pp. 594. £50 (hbk); £19.95 (pbk).

THE urge to try to answer the questions 'where', 'when' and 'why' in connection with the acquisition of any human attribute, whether behavioural or morphological, evidently holds a special fascination. It is a measure of the power of this fascination that scholars continue to embark upon quests for 'origins' despite all the accumulated wisdom that suggests that many of these projects are destined to end in disappointment.

This dismal prognosis is due to the combination of the adage 'absence of evidence is not evidence of absence' and the inevitability that the first good (that is, thoroughly convincing) evidence of a behaviour or morphology almost certainly significantly postdates its initial adoption or appearance. There is also the added complication that many quests for origins are bedevilled by an interpretation of history that does not countenance any other state for the attribute or behaviour than the prevailing one. To be scored as 'present', a behaviour or morphology has to be fully fledged; partially developed states do not count in such schemes.

Human evolution has its fair share of origins conundrums. Modern humans and chimpanzees are very obviously different, but when and where in the fossil record do we find the first convincing evidence of precursors that were the exclusive ancestors of modern humans? Early hominid species still have an apelike body-shape even though they were capable of bipedal locomotion: when and where did the first hominid with a more human-like body size and shape appear? When and where did hominids first manufacture and use tools? When and where did creatures indistinguishable from modern humans appear in the fossil record?

At or near the top of most 'wish-lists' of what people would dearly like to know about the evolutionary history of modern human behaviour would be when, where and why were agriculture and pastoralism established as subsistence strategies? Of all human innovations, agriculture has been one of the most important. The ability to coax higher and higher yields from agricultural enterprises on all scales has allowed the modern human population to expand to an extent that would be unthinkable in other circumstances. Was

agriculture introduced incrementally, or as an all-or-nothing package? Did it appear independently in several regions across the Old World, or was it such a potent innovation that, once discovered, it spread rapidly, much like a bush-fire? Were agricultural skills transmitted by example, or by the migration of the people who practised them? In this volume, David Harris has brought together the most recent evidence germane to understanding how agriculture came to play this pivotal role.

Harris is careful to draw a distinction between cultivation and agriculture — the latter being distinguished by the dominance of domesticated plants and by the appearance of domesticated animals. This is a convention that most of the 27 contributors are evidently happy to accept. There is also widespread support among the contributors for the 'conventional wisdom' about the origin of agriculture, which holds that in Eurasia there were only two main foci for agricultural innovations. One, in southwest Asia in the region that has become known as the 'Fertile Crescent', was a development based on cereals and pulses and on the domestication of goats and sheep. The other, in China in the Huanghe and the Yangzi basins, was based on millets and rice and involved the domestication of pigs and chickens.

Despite substantial differences in the kinds of plants and animals that were domesticated, and the several thousands of kilometres that separated the two centres of agricultural innovation, the effective coincidence in the timing of the two processes — both occurred in, or around, the eighth and the seventh millennia — suggests to many of the contributors that the adoption of agriculture was linked with the global deterioration in climate known as the 'Younger Dryas'. This resulted in colder, wetter winters and hotter, drier summers. The effect in the Levant was to exacerbate the spread of oak-dominated woodland and to amplify a corresponding increase in carrying capacity, especially of annual grasses. The rest, as they say, is history.

Although many of the contributions add their support to the 'two foci' model, relatively few of the contributors provide criteria for deciding whether the evidence for agriculture in the form of domesticated plants and animals is sufficiently strong. Without being provided with explicit criteria, it is often difficult for an outsider to judge whether the presence of agriculture is being diagnosed on the basis of primary evidence, such as plant morphology, or on the basis of secondary or even tertiary inference, such as the use of pottery.

Colin Renfrew, for example, suggests

that his linguistic data provide compelling circumstantial evidence for the hypothesis that the spread of agriculture and language followed a similar pattern. The congruence between the inferred pattern of the spread of language and the genetic evidence for diffusion is indeed a close one, but we need to remember that these categories of evidence are still circumstantial.

All this agreement with conventional wisdom makes me uneasy. It is as well to remember that C. H. Waddington, while he referred apparently deferentially to the "conventional wisdom of the dominant group", also encouraged the use of the less-than-deferential euphemism 'coudung' for the condition in which there is widespread satisfaction that all the evidence points in the same direction. There are uncanny resemblances between the scenario for the origin of agriculture outlined in this volume and suggestions that have been made for the origin of two distinctive groups of fossil hominids, those belonging to the genera *Homo* and *Paranthropus*.

However, instead of the Younger Dryas providing the substrate for the adoption of agriculture, the drying and cooling between 3 and 2 million years ago have been implicated in promoting an adaptive dichotomy in the early hominid, with *Paranthropus* turning to megadonty (that is, enlarging the chewing capacity of its teeth) and *Homo* to culture. Since the apparent link with climate has proved difficult, if not impossible, to establish in the hominid case, similarly the Younger Dryas scenario for agriculture strikes me as too simplistic.

The breadth of evidence being brought to bear in the pursuit of agricultural origins is as impressive as the ingenuity with which much of it is being applied. The contributors in this volume have evidently been chosen with care and the editing is exemplary, as are the summary chapters written by the editor. The contributions to an equivalent book in a decade's time will doubtless give more prominence to molecular evidence, and I would hope that the criteria for recognizing agriculture will have been developed further and that more work will have been done in Africa.

Lastly, I hope that the archaeologists of the future will choose a more reductionist approach. There is more than a hint in some of the contributions that the 'present' is being imposed on the 'past'. Our ancestors may have been less certain about the distinctions between foraging, cultivation, herding and agriculture than we are today. □

Bernard Wood is in the Department of Human Anatomy and Cell Biology, The University of Liverpool, Liverpool L69 3BX, UK.

Reading the mind of God

Maryanne Traylen

Science and Wonders: Conversations about Science and Belief. By Russell Stannard. *Faber and Faber*: 1996. Pp. 208. £8.99 (pbk).

IN his book *Understanding the Present: Science and the Soul of Modern Man*, Brian Appleyard asks whether science is incompatible with the word of God. By arguing that science cannot be value-free or neutral — it is, he claims, spiritually corrosive — he encourages a division between science and the spiritual. In *The Unnatural Nature of Science*, Lewis Wolpert encourages the same division from a converse angle. In the face of science, he claims, the spiritual is non-existent. But in *The Mind of God*, Paul Davies, concerned to put the mind back into matter, tries to show how science should be enriching rather than alienating.

Russell Stannard's gentle enquiry, which arose from a series of BBC radio programmes, is not as deliberate in its intent as any of these recent books. He does, however, interview those who, like himself, are in agreement with Davies

and want to see the gap created by dualism bridged. The people he has talked to are theologians as well as former physicists (Stannard is himself a professor of physics) or biochemists, or are involved as directors or lecturers in theology or biblical studies as well as science. He also interviews atheists, astronomers, neuroscientists, biologists and even creationists — all in the cause of strengthening the interface between science and religion, which, he claims, is one of the fastest growing areas of academic study today. Amiable discussion flows from the questions put to his interviewees about the cosmos, life, the mind, room for God and the relationship between science and religion. Yet the overall result, after these big issues have been chattily chewed over, is unsatisfactorily short of target.

Stannard lets the views of others paint a picture, be they the Archbishop of York in his palace, the psychiatrist Montague Barker on Freud's couch in Hampstead or Jocelyn Bell-Burnell, for a while the only British woman professor of physics, all of whom discuss whether the uncertainty principle and the Big Bang theory imply the existence of a Creator. Keith Ward attempts to remove the crisis from this dilemma by substituting the notion of the 'creation' of the Universe — which takes place not only at the beginning, but also now and always — for the notion of



SQUADRON of squid — a group of *Sepioteuthis sepiodea* in normal daytime resting position above a coral reef at Little Cayman Island. Cephalopods, of which there are about 700 species (including squids, octopuses and cuttlefishes) living throughout the world's oceans, are a very ancient and specialized group within the Mollusca. They are considered to be the most highly evolved marine invertebrates, with elaborate sense organs, large brains and complex behaviour. From *Cephalopod Behaviour* by Roger T. Hanlon and John B. Messenger. Cambridge University Press, £50.

the 'origin'. An agnostic scientist talks about the way in which science reclaims land for a little island of knowledge from a huge sea of ignorance. A biologist talks of evolution, natural selection and the possibility Darwin affords us of God being *in* life, being immanent as well as transcendent.

Serious issues raised include the idea of interplay between chance and natural law, of evolution as helpful in the development of new creative things, rather than acting against religion; Will Provine's controversial idea of reprogramming people; the Human Genome Project as the Mount Everest of scientific enquiry; Jung's belief that Christianity has distanced spirit and matter, and the importance of integrating matter into the spiritual world; the ability of science to resolve what is meant by consciousness and to solve 'everything'; Stephen Hawking's reasoning that without time or space before the Big Bang, God could not have existed — refuted by Augustinian thought showing God to be not just in space-time, but also in and beyond; Hell as the absence of God; religion as a virus that spreads; science as God; science as a path to God; and the importance of combining science and religion in education.

It is a rich assortment of essentials all touched upon: varied and interesting points of view are brought up as quickly as they are passed over. Stannard does not make a strong case for the dialogue he wants between the opposing disciplines; he could for instance have made his achievement more emphatic by stressing in his own debate William Blake's maxim: "Without contraries is no progression".

He has put the evidence before us, with only polite indignations showing us what offends. But without a central thread, he leaves himself open to the accusations of those who feel that because science and theology are separate disciplines anyway, providing meaning through different channels, there is really no argument in the first place.

Stannard's passion may not always be apparent to the reader: his is a difficult task of incorporating everyone's views into a cohesive whole, and the result is frustratingly broad rather than deep; there is nothing for us to get our teeth into. Yet he cannot be wrong in advocating that science and theology should be approached with 'wonder'. He can only be too safe. Where is the wonder for science or belief when it is a word used at the end rather than at the beginning of the book? If Stannard is not clear about the exact connection between science and belief, he has at least established a dialogue and shown that it is entering society's consciousness. □

Maryanne Tryalen is at 8 Dashwood Road, London N8 9AD, UK.

Yesterday's weather

Peter Lynch

Calculating the Weather: Meteorology in the 20th Century. By Frederik Nebeker. Academic: 1995. Pp. 255. £44, \$64.95.

How did meteorology join the ranks of the exact sciences, and how is weather prediction done today? With few general accounts of computer forecasting available, the appearance of Frederik Nebeker's book aroused hopes that here at last was a comprehensive discussion of these questions. But disappointment awaits the reader seeking an understanding of modern weather forecasting. Nebeker looks at the ever-increasing role of calculation in meteorology but fails to provide a satisfactory picture of the state of play.

The central theme of the book is that computational methods have brought about a unification of meteorology. In 1900, three main strands were clearly discernible: climatology, dynamical theory and practical forecasting. Over the course of this century, these three strands have become more quantitative and computational and a greater unity has become apparent. To the unifying forces acting in the three areas, Nebeker gives the self-explanatory if prosaic names 'data-push', 'theory-pull' and 'science-not-art'.

Nebeker starts by reviewing the three strands as they were in 1900. The then status of climatology is captured in a quotation from Joseph Henry: "There is perhaps no branch of science relative to which so many observations have been made and so many records accumulated, and yet from which so few general principles have been deduced." The theoretical tradition is reviewed but the treatment is patchy. The dynamical studies of Ferrel, Hann and Exner are outlined and there is a mini-essay on Sir Napier Shaw, but one is left without any overall picture of the status of theoretical meteorology in 1900. Things improve with the chapter on practical prediction. The synoptic method is well described, giving a good feeling for the plight of the hapless forecaster.

In the second part, Nebeker deals with developments up to 1945. Bjerknes's programme to make forecasting a science is described, although the crucial 1904 reference to his plan is missing. The graphical solution method is well illustrated by an example of calculating vertical velocity near the ground. Richardson's method and forecast are the subject of a long and interesting chapter. Nebeker states that Richardson's work was highly regarded, but he does not adequately explain why no-one followed it up, even though Richardson had pointed to smoothing as

the solution to his problems.

The growth of meteorology between the two world wars is considered next, especially the dramatic influence of frontal theory on practical forecasting. Nebeker makes the dubious claim that the Bergen school methods were "probably more subjective than earlier procedures". There is a brief but interesting biographical sketch of Rossby. The remainder of this part concerns calculating aids used before and during the Second World War. The extensive treatment of punched-card equipment is rather dull, but the descriptions of the D-day forecasts and of the wartime importance of meteorology are good.

In the third part, Nebeker considers the emergence of computer forecasting, providing excellent coverage of the Princeton project of von Neumann and the crucial role played by Jule Charney. The subsequent introduction of computational methods into general operations and the greater cohesiveness that has resulted are then discussed, but with virtually no information on work since 1970. For example, the European Centre for Medium-Range Weather Forecasts is not even mentioned.

The final chapter looks at Lorenz's discovery of chaos, which has brought about a scientific revolution and has profound implications for predictability of weather. A major thrust of recent research has been to estimate the probability of different weather scenarios using ensembles of forecasts. This springs directly from Lorenz's work on sensitivity to initial conditions. Yet ensemble forecasting, which is inextricably intertwined with supercomputing, is not mentioned; such an omission in a book of this nature is extraordinary.

The book is well written and pleasant to read and includes about 450 references and a good index, although the end-notes beloved of publishers are a nuisance for readers. There are several errors in the few equations displayed and even the list of equations on p. 51 is wrong. But my main complaint is that the book might have been written 30 years ago. Readers interested in the early history of numerical forecasting will find it a useful summary. Given the title, most scientists will be unhappy with the absence of a description of modern forecasting methods. ▲

Peter Lynch is at Met Eireann, Glasnevin Hill, Dublin 9, Ireland.

New in paperback

The Second Creation: Makers of the Revolution in Twentieth-Century Physics by Robert P. Crease and Charles C. Mann. Rutgers University Press, \$21.95. "...a satisfying popularization of what may be the most glamorous and philosophically nourishing part of this century's physics", said Laurie M. Brown in his review of the original edition (*Nature* **323**, 119; 1986).

A search for human influences on the thermal structure of the atmosphere

B. D. Santer^{*}, K. E. Taylor^{*†}, T. M. L. Wigley[‡], T. C. Johns[§], P. D. Jones^{||},
D. J. Karoly[¶], J. F. B. Mitchell[§], A. H. Oort[#], J. E. Penner[†], V. Ramaswamy[#],
M. D. Schwarzkopf[#], R. J. Stouffer[#] & S. Tett[§]

^{*} Program for Climate Model Diagnosis and Intercomparison, [†] Atmospheric Science Division, Lawrence Livermore National Laboratory, Livermore, California 94550, USA

[‡] National Center for Atmospheric Research, Boulder, Colorado 80307-3000, USA

[§] Hadley Centre for Climate Prediction and Research, Meteorological Office, Bracknell RG12 2SY, UK

^{||} Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, UK

[¶] Cooperative Research Centre for Southern Hemisphere Meteorology, Monash University, Clayton VIC 3168, Australia

[#] NOAA/Geophysical Fluid Dynamics Laboratory, PO Box 308, Princeton University, Princeton, New Jersey 08542, USA

The observed spatial patterns of temperature change in the free atmosphere from 1963 to 1987 are similar to those predicted by state-of-the-art climate models incorporating various combinations of changes in carbon dioxide, anthropogenic sulphate aerosol and stratospheric ozone concentrations. The degree of pattern similarity between models and observations increases through this period. It is likely that this trend is partially due to human activities, although many uncertainties remain, particularly relating to estimates of natural variability.

CHANGES in the vertical structure of atmospheric temperature have been proposed as a possible 'fingerprint' of greenhouse-gas-induced climate change¹⁻⁴. Until recently, most of the information about such a fingerprint resulted from experiments in which an atmospheric general circulation model (AGCM) coupled to a mixed-layer ocean was forced by a doubling of atmospheric CO₂ levels.^{5,6} For annually averaged changes, these experiments show a hemispherically symmetrical pattern of stratospheric cooling and tropospheric warming, with a warming maximum in the tropical upper troposphere (Fig. 1a). One recent study⁷ compared such model-predicted signals with observed radiosonde measurements⁸ and concluded that this fingerprint was increasingly evident in observed data. The degree of time-increasing similarity was judged to be statistically significant—that is, unlikely to be due to natural climate variability alone.

The work reported here differs from this earlier work in four respects. First, we examine the relative detectability of vertical temperature-change signals due to individual and combined changes in atmospheric CO₂ and anthropogenic sulphate aerosols^{9,10}. Previous detection work involving temperature changes in the free atmosphere has used signals due to increases in CO₂ only. We consider signals due to combined CO₂ and aerosol forcing because recent studies of near-surface temperature changes show that a combined signal may be easier to identify in the observations than a signal due to changes in CO₂ alone^{10,11}.

Second, we use signal data from two different models in order to explore the sensitivity of our results to model-dependent uncertainties in the definition of an anthropogenic signal. These uncertainties arise from model differences in physics, resolution, representation of aerosol effects, and modelling strategy.

Third, we consider how a combined CO₂ + aerosol vertical temperature-change signal might be modified by observed changes in stratospheric ozone. The observed reduction in stratospheric ozone over the past two decades is attributable largely to the industrial production of halocarbons¹². One recent study in which an AGCM was forced by changes in both CO₂ and stratospheric ozone showed that the inclusion of ozone effects improves model agreement with the radiosonde temperature data, particu-

larly in the upper troposphere^{13,14}. As results are not yet available from experiments with combined changes in CO₂, anthropogenic sulphate aerosols and stratospheric ozone, we perform a simple sensitivity study of possible ozone effects by linearly combining results from CO₂ + aerosol⁹ and ozone-only¹⁵ model studies.

Fourth, we use information from three long control runs performed with coupled atmosphere–ocean GCMs (CGCMs) to assess the significance of trends in model-versus-observed pattern similarity. Such integrations provide estimates of internally generated natural climate variability on a range of space and time-scales. Comparable information on the multi-decadal natural climate variability crucial to the detection problem is impossible to obtain from the short (<40-year) radiosonde temperature record.

Model signals and observational data

We use model data from experiments with fixed levels of anthropogenic pollutants ('equilibrium experiments') and from integrations in which the atmospheric concentrations of greenhouse gases and aerosols increase over time ('transient experiments'). The equilibrium CO₂-only and CO₂ + aerosol vertical temperature-change signals were taken from experiments performed by Taylor and Penner⁹ (henceforth TP) with a tropospheric chemistry model^{16,17} coupled to an AGCM with a simple mixed-layer ocean¹⁸. The chemistry–climate model considers only the direct radiative effects of sulphate aerosols (reflection of incident solar radiation)¹⁹. In addition to a control run with nominal pre-industrial levels of CO₂ (270 parts per million by volume, p.p.m.v.) and no anthropogenic sulphur emissions, three perturbation experiments were performed: a sulphate-only experiment (S-TP) with near-present-day anthropogenic sulphur emissions, a CO₂-only experiment (C-TP) with near-present-day CO₂ levels (345 p.p.m.v.), and an experiment with combined present-day CO₂ levels and anthropogenic sulphur emissions (SC-TP)⁹. The differences between the perturbation experiments and the control represent equilibrium changes from pre-industrial to present-day conditions.

We also consider vertical temperature-change signals from two

recent transient experiments performed with the Hadley Centre (henceforth HC) CGCM^{10,20}. The model has full ocean dynamics, and was forced over 1860 to 1990 with historical increases in greenhouse gases only (C-HC) and with increases in both greenhouse gases and anthropogenic SO₂ emissions (SC-HC). After 1990 the concentration of SO₂ evolved according to IPCC Scenario IS92a²¹ while the concentration of equivalent CO₂ increased at 1% per year. Direct scattering effects of aerosols are represented by changes in the surface albedo²².

To study how a reduction in stratospheric O₃ might modify the SC signal pattern, we use data from an experiment performed with the Geophysical Fluid Dynamics Laboratory (GFDL) 'SKYHI' atmospheric GCM¹⁵. The model was forced with observed monthly-mean zonal average changes in stratospheric ozone over the period 1979–90, and was run with fixed cloud distributions in the troposphere and sea-surface temperature prescribed according to climatology. An idealized vertical structure of ozone losses was imposed, with constant percentage reductions in an atmospheric region extending from the tropopause to roughly 7 km above¹⁵.

The radiosonde temperature analyses were available as anomalies for December–January–February (DJF), June–July–August (JJA) and annually averaged data relative to a reference period of 1963–73, and spanned the period 1963–87. The data are in the form of zonal averages for seven atmospheric levels (850, 700, 500, 300, 200, 100 and 50 hPa). The principal data uncertainties have been described previously^{7,8,23}.

Comparisons between the radiosonde data and satellite-derived estimates of vertical temperature changes indicate that the two data sets are in good agreement over the period of overlap, at least in terms of the global²⁴ and hemispheric⁸ means. For the present study we compared the radiosonde data with a reanalysis of operationally produced climate data²⁵. The seasonal patterns of (zonally averaged) linear trends as a function of latitude and height were in close agreement for the period of overlap between the two data sets (1979–87), despite differences in the spatial coverage of the radiosonde data (primarily land only) and the reanalysis (land + ocean).

Note that the amplitudes of the observed changes and model signals are not directly comparable, as they represent responses to radiative forcing changes over different periods. If the radiative forcing histories and lags between forcing and response were known exactly for O₃, CO₂ and sulphate aerosols, it would be possible to make a more meaningful comparison of the amplitudes of observed and modelled vertical temperature changes by scaling according to differences in overall response. Large forcing uncertainties, particularly for sulphate aerosol effects, make such scaling exercises very difficult. We circumvent some of these difficulties by using a comparison statistic that focuses on patterns rather than amplitudes of temperature change. The issue of amplitude differences is important in the linear superposition of O₃ and SC signals, and we return to it later.

Patterns of vertical temperature change

Modelled and observed patterns of annual-mean zonal-mean temperature change as a function of latitude and height are shown in Fig. 1. The pattern of stratospheric cooling and tropospheric warming in C-TP (Fig. 1a) is in accord with previous modelling work^{5–7} and represents the direct radiative signature of the change in CO₂. Maximum warming occurs in the tropical upper troposphere, and temperature changes are hemispherically symmetric. In contrast, both the S-TP (Fig. 1b) and SC-TP (Fig. 1c) signals show a hemispherically asymmetric response, with (respectively) increased cooling and reduced warming in the Northern Hemisphere, where anthropogenic sulphate aerosol forcing is largest^{9,22}. The SC-TP integration (like C-TP) also shows the dipole structure of stratospheric cooling/tropospheric warming characteristic of CO₂ changes, whereas there is minimal stratospheric response in S-TP.

The C-HC and SC-HC transient results (Fig. 1d, e) have

FIG. 1 Modelled and observed zonal-mean annually averaged changes (°C) in the thermal structure of the atmosphere. The equilibrium experiments by Taylor and Penner (TP)⁹ simulate temperature changes for nominal 'present-day' levels of atmospheric CO₂ only (C-TP; a), anthropogenic sulphate aerosols only (S-TP; b), and combined forcing by CO₂ + sulphate aerosols (SC-TP; c) relative to a control run with pre-industrial levels of CO₂ and no anthropogenic sulphur emissions. All TP integrations were at least 30 years in duration, and temperature-change signals were computed using averages over the last 20 years of the control run and each perturbation experiment. Patterns of the response to time-varying increases in greenhouse gases only (C-HC; d) and in greenhouse gases and aerosols (SC-HC; e) were taken from simulations performed with the Hadley Centre CGCM^{10,20}. Temperature-change signals are the decadal averages of C-HC and SC-HC for the modelled '1990s' expressed relative to the respective C-HC and SC-HC averages over 1880–1920. The possible effects of stratospheric ozone reduction over the period 1979–90 (f) are from a recent equilibrium experiment by Ramaswamy *et al.*¹⁵ The sensitivity studies COMB1 (SC-TP + O₃; g) and COMB2 ($\frac{1}{2}$ SC-TP + O₃; h) consider the possible effects of stratospheric O₃ depletion on the SC-TP signal. COMB3 ($\frac{1}{2}$ S-TP + C-TP; i) illustrates the sensitivity of model-observed pattern similarities to a possible overestimate of direct aerosol effects in TP. Observed changes (j) are radiosonde-based temperature measurements from the data set by Oort⁸, and are expressed as total least-squares linear trends (°C) over the 25-year period extending from May 1963 to April 1988.

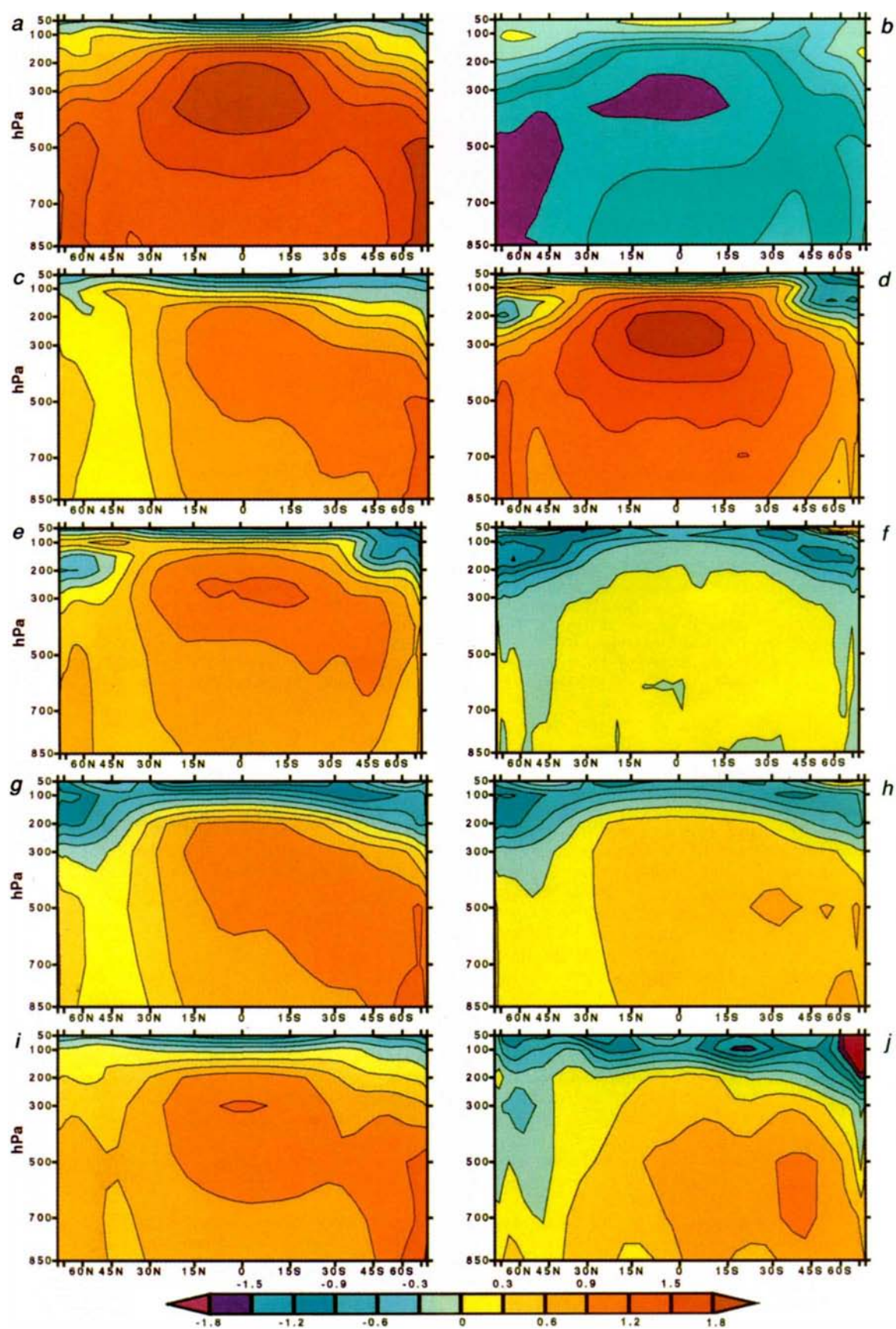
features which are qualitatively similar to the equilibrium response patterns from the corresponding TP experiments (Fig. 1a, c). Both types of experiment yield stratospheric cooling and tropospheric warming (C and SC) and reduced warming in the Northern Hemisphere (SC only). We return to this point later, because it is relevant to the issue of the usefulness of equilibrium signals in climate-change detection studies.

Vertical temperature changes due to stratospheric O₃ reduction (Fig. 1f) are characterized by stratospheric cooling, with maximum cooling at high latitudes in both hemispheres. Owing to dynamical effects, cooling occurs throughout the lower stratosphere, even at low latitudes where the imposed ozone changes are negligible^{15,26}. The response is not hemispherically symmetric: stratospheric cooling that is statistically significant¹⁵ occurs over a wider latitude range in the Northern Hemisphere than in the Southern Hemisphere. This is primarily due to a hemispheric asymmetry in the observed ozone changes. Some of the model-observed temperature differences in Fig. 1f and j, such as the warming above about 70 hPa poleward of 45°S, are probably related to the idealized altitudinal profile of ozone loss²⁶. Other differences are due to the different time periods considered in the model experiment and in the observations.

To examine the possible effects of stratospheric O₃ depletion we perform two sensitivity studies. COMB1 is the unweighted sum of the SC-TP and O₃ signals (Fig. 1g). COMB2 (Fig. 1h) explores uncertainties in the relative amplitudes of the SC-TP and O₃ signal components by halving the amplitude of the SC-TP signal. A third sensitivity study (COMB3; Fig. 1i), considers the effect of uncertainties in the magnitude of the aerosol forcing (and response) by halving the amplitude of the S-TP signal relative to the C-TP signal.

For any of these sensitivity experiments to be a realistic estimate of the response to combined CO₂ + SO₄ + O₃ forcing requires the relative weights of the individual forcings to be realistic and the climate system to respond quasi-linearly to perturbations. We have tested this linear superposition assumption and found it to be valid for combining the C-TP and S-TP signals²⁷. The assumption cannot be tested for the HC experiments (owing to the lack of a transient experiment with forcing by aerosol effects alone) or for O₃ effects (because suitable model studies with individual and combined forcing by CO₂ + SO₄ + O₃ were not available). A linear combination of O₃ and SC effects is probably reasonable above the tropopause²⁸.

Incorporating stratospheric ozone effects (COMB1, COMB2) intensifies stratospheric cooling and reduces the height of the



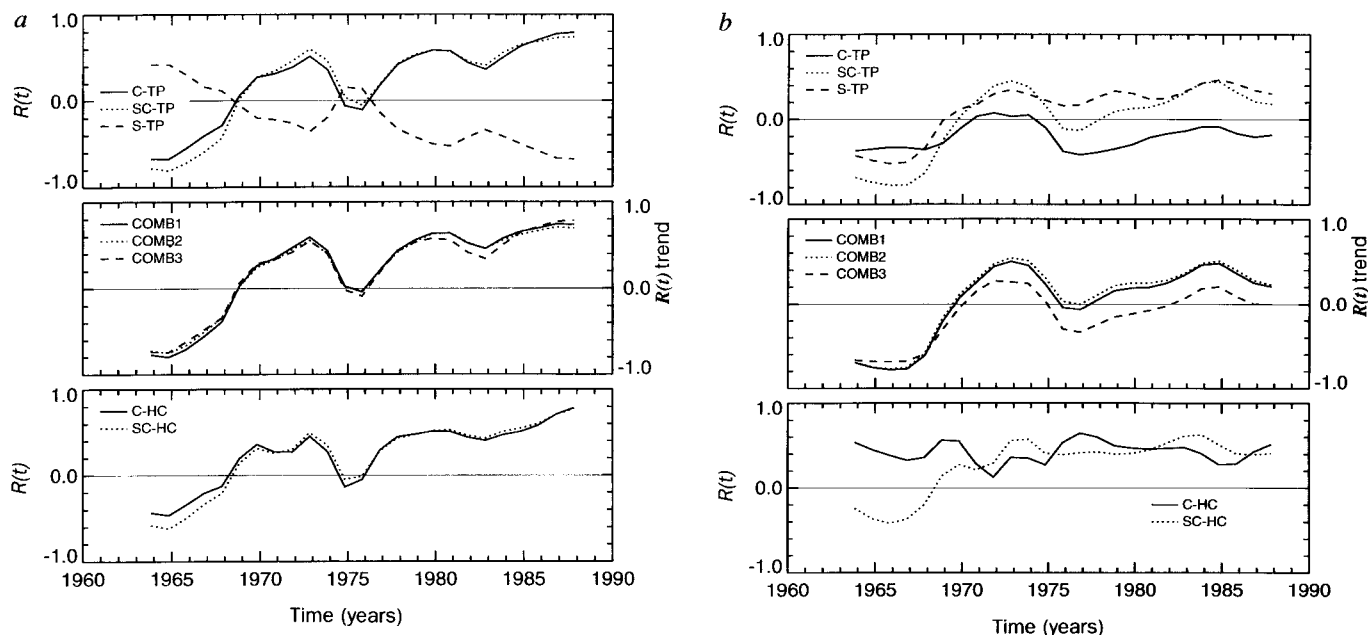


FIG. 2 Time series of centred pattern correlations, $R(t)$, between model-predicted and observed changes in zonal-mean latitude–height profiles of atmospheric temperature. For sources of model and observed data, refer to Fig. 1. Observed changes are expressed as a sequence of anomaly patterns spanning the period 1963–87, and were smoothed with a 5-term binomial filter to suppress variability on timescales shorter than a decade (for example, associated with ENSO events and the quasi-biennial oscillation)¹¹. For each model experiment, one pattern characterizes the response to the imposed anthropogenic forcing (see Fig. 1). This fixed pattern is correlated with the observed time-varying spatial patterns. Results are for temperature-change patterns defined over the full vertical extent of the radiosonde data (50–850 hPa; a) and over the mid-

to lower-troposphere only (500–850 hPa; b). The premise underlying the use of a centred correlation for attribution is that different ‘causes’ (forcing mechanisms) have different response patterns. If one can demonstrate time-increasing correspondence between the observations and a model-predicted pattern of change, and show that correspondence exists at hemispheric or smaller spatial scales—not only at the surface, but also in the full three-dimensional structure of the atmosphere—then it is less likely that forcing mechanisms other than the ones being considered could match the predicted response pattern. The pattern correlations shown here were computed using annual averages and with pressure- and area-weighted data. Trends in $R(t)$ were relatively insensitive to the choice of averaging period for defining observed anomalies¹¹.

transition between stratospheric cooling and tropospheric warming¹⁴. However, the transition height in COMB1 and COMB2 is still roughly 50 hPa higher than in observations (compare Fig. 1g, h and Fig. 1j). There are several possible reasons for this. These include: uncertainties in observed O_3 losses and thus in the simulated temperature changes in the vicinity of the tropopause^{15,26}; concerns regarding the validity of the linearity assumption for combining SC-TP and O_3 signals; and the (neglected) effects of upper tropospheric O_3 changes. The coarse vertical resolution of the observed data and most of the model signals also hampers more accurate determination of discrepancies in transition height.

The magnitude of direct anthropogenic sulphate aerosol forcing in TP (roughly -0.9 W m^{-2}) is just above the upper end of the range estimated in recent reviews^{29,30}, although clearly within the range given for total (direct + indirect) sulphate aerosol forcing. It is relevant in this regard that the hemispheric-scale patterns of forcing due to aerosol direct and indirect effects may be similar^{31–34}. The effect in COMB3 of halving the temperature response in S-TP relative to C-TP is to reduce the interhemispheric asymmetry in the low- to mid-troposphere and enhance the temperature-change contrast between the stratosphere and troposphere (Fig. 1i).

Figure 1j shows observed temperature changes, expressed as linear trends over 1963–87. The observations show two prominent features: stratospheric cooling and tropospheric warming, and reduced warming in the Northern Hemisphere between 850 and 300 hPa. These have been documented in previous investigations^{3,4,7,35}. The pattern of stratospheric cooling and tropospheric warming is common to the observations and all model signals shown in Fig. 1 (with the exception of the ‘ozone-only’ and S-TP signals). In the lower atmosphere, it is clear that

the observations are in better accord with combined greenhouse-gas and aerosol signals than with ‘ CO_2 -only’ signals^{10,11}.

Model–observed pattern similarity

To compare model and observed vertical temperature-change patterns we use a so-called ‘centred’ correlation statistic¹¹, $R(t)$:

$$R(t) = \left[\sum_{x=1}^n (\Delta D(x, t) - \widehat{\Delta D}(t)) (\Delta M(x) - \widehat{\Delta M}) \right] / \left[n s_D(t) s_M \right] \quad (1)$$

ΔD and ΔM are temperature-change fields for observed data and model output, respectively, and x and t are discrete indices over space ($x = 1, \dots, n$, the combined latitude–height dimension of the radiosonde data) and time ($t = 1963, \dots, 1987$). Observed changes are anomalies relative to the average over 1963–73, and model changes are differences between time averages of perturbation and control experiments (see Fig. 1). The $\widehat{\cdot}$ indicates a spatial average. The observed spatial variance $s_D^2(t)$ is

$$s_D^2(t) = \sum_{x=1}^n \left[\Delta D(x, t) - \widehat{\Delta D}(t) \right]^2 / (n - 1) \quad (2)$$

with the model spatial variance s_M^2 defined similarly.

If the observed time-varying patterns of temperature change are becoming increasingly similar to the model-predicted responses, the $R(t)$ statistic will show a sustained positive trend³⁶. This trend is unlikely to be linear and monotonic, as the observations reflect a response to the real (as opposed to the modelled) anthropogenic forcing and to other human-induced and natural forcings not included in the model experiments. Additionally, any overall $R(t)$ trend will be modulated by internally

TABLE 1 Significance test results

		Trend lengths (years)								
		10			15			25		
Season	Signal	MPI	GFDL	HC	MPI	GFDL	HC	MPI	GFDL	HC
a 50–850 hPa										
DJF	C-TP	0.01	0.00	0.05	0.00	0.00	0.01	0.00	0.00	0.00
	C-HC	0.06	0.03	0.21	0.02	0.00	0.09	0.01	0.00	0.02
	SC-TP	0.02	0.02	0.12	0.00	0.00	0.02	0.00	0.00	0.00
	SC-HC	0.07	0.02	0.23	0.02	0.00	0.11	0.00	0.00	0.01
	COMB1	0.02	0.04	0.05	0.00	0.00	0.00	0.00	0.00	0.00
	COMB2	0.02	0.05	0.03	0.00	0.00	0.00	0.00	0.00	0.00
	COMB3	0.01	0.00	0.05	0.00	0.00	0.01	0.00	0.00	0.00
JJA	C-TP	0.08	0.20	0.25	0.02	0.08	0.16	0.00	0.00	0.00
	C-HC	0.23	0.35	0.39	0.07	0.17	0.22	0.00	0.00	0.00
	SC-TP	0.10	0.17	0.24	0.07	0.10	0.20	0.00	0.00	0.00
	SC-HC	0.13	0.25	0.32	0.02	0.10	0.17	0.00	0.00	0.00
	COMB1	0.10	0.18	0.20	0.06	0.10	0.14	0.00	0.00	0.00
	COMB2	0.10	0.20	0.20	0.05	0.11	0.14	0.00	0.00	0.00
	COMB3	0.10	0.19	0.24	0.05	0.09	0.15	0.00	0.00	0.00
ANN	C-TP	0.16	0.24	0.33	0.01	0.01	0.11	0.00	0.00	0.00
	C-HC	0.20	0.26	0.36	0.04	0.04	0.19	0.00	0.00	0.01
	SC-TP	0.23	0.27	0.38	0.04	0.03	0.15	0.00	0.00	0.00
	SC-HC	0.20	0.26	0.36	0.05	0.04	0.18	0.00	0.00	0.00
	COMB1	0.26	0.34	0.38	0.04	0.04	0.10	0.00	0.00	0.00
	COMB2	0.30	0.38	0.38	0.03	0.05	0.07	0.00	0.00	0.00
	COMB3	0.15	0.21	0.33	0.02	0.01	0.13	0.00	0.00	0.00
b 500–850 hPa										
DJF	C-TP	0.08	0.15	0.20	0.06	0.12	0.15	0.08	0.12	0.25
	C-HC	0.55	0.53	0.53	0.49	0.51	0.48	0.20	0.28	0.27
	SC-TP	0.29	0.37	0.40	0.30	0.38	0.36	0.01	0.04	0.02
	SC-HC	0.48	0.50	0.47	0.51	0.52	0.50	0.00	0.02	0.01
	COMB1	0.20	0.29	0.29	0.32	0.37	0.36	0.00	0.03	0.01
	COMB2	0.16	0.26	0.21	0.33	0.39	0.37	0.00	0.02	0.01
	COMB3	0.10	0.17	0.19	0.12	0.21	0.18	0.07	0.11	0.09
JJA	C-TP	0.34	0.46	0.43	0.45	0.48	0.43	0.55	0.51	0.44
	C-HC	0.81	0.66	0.63	0.50	0.52	0.46	0.53	0.51	0.52
	SC-TP	0.17	0.28	0.31	0.18	0.29	0.28	0.02	0.11	0.10
	SC-HC	0.20	0.31	0.31	0.22	0.33	0.31	0.01	0.04	0.06
	COMB1	0.18	0.28	0.30	0.20	0.29	0.28	0.02	0.11	0.09
	COMB2	0.20	0.28	0.29	0.21	0.31	0.28	0.03	0.12	0.09
	COMB3	0.19	0.33	0.36	0.21	0.33	0.33	0.07	0.16	0.13
ANN	C-TP	0.24	0.37	0.36	0.32	0.43	0.41	0.27	0.40	0.37
	C-HC	0.66	0.57	0.54	0.61	0.55	0.52	0.44	0.47	0.43
	SC-TP	0.29	0.41	0.38	0.22	0.31	0.31	0.00	0.03	0.04
	SC-HC	0.54	0.49	0.48	0.47	0.50	0.49	0.00	0.02	0.02
	COMB1	0.33	0.42	0.40	0.27	0.35	0.34	0.00	0.02	0.03
	COMB2	0.37	0.44	0.41	0.32	0.38	0.37	0.00	0.02	0.00
	COMB3	0.19	0.35	0.34	0.21	0.34	0.33	0.02	0.08	0.06

Significance levels (p -values) for seasonal and annual vertical temperature-change signals from the TP and HC experiments with forcing by CO₂ only and CO₂ + aerosols and from the three sensitivity studies (COMB1, COMB2, COMB3). The first column gives the season of interest, while the second column indicates the experiment from which the signal pattern was taken. The remaining columns give the model experiments used to obtain natural variability estimates (on three different timescales). The signals of interest are the linear trends for the most recent 10, 15 and 25 years of the $R(t)$ time series shown in Fig. 2a and b—that is, the trends over 1978–87, 1973–87, and 1963–87. These trends provide information on the degree of time-increasing pattern similarity between the observations and the model simulations. Distributions of ‘unforced’ $R(t)$ trends on 10-, 15- and 25-year timescales were generated as described in Fig. 3. Significance levels were then computed by comparing the ‘signal’ $R(t)$ trends with the appropriate sampling distributions for unforced trends¹¹. Shaded boxes denote results that achieve significance at the 5% level or better. In these cases, the time-increasing similarity between model signal patterns and observations is unlikely to be due to (model estimated) internally-generated natural variability. Results are for model–observed comparisons over 50–850 hPa (a) and over 500–850 hPa (b).

generated natural climate variability. There are two main issues of interest: whether trends in $R(t)$ could be due to natural internal variability alone, and whether they are different for different model signals.

For comparisons over 50–850 hPa (Fig. 2a), the behaviour of $R(t)$ is similar for all signals except S-TP: trends are positive over the full 25-year period, indicating an increasing expression of the model-predicted patterns in the observed annually averaged data. These similarities are due partly to the large-amplitude pattern of

stratospheric cooling/tropospheric warming common to all signals that incorporate CO₂ effects (see Fig. 1). Differences in the magnitude of this common pattern in the individual signals are reduced since they are scaled by different model spatial variances, s_M^2 . Similarities in $R(t)$ are also related to the removal of different spatial mean values (ΔM) from the model signal patterns, which has the effect of making the height of the transition from stratospheric cooling to tropospheric warming more similar in the individual signals.

To focus on hemispheric asymmetry differences in the various signals, we then restricted pattern comparisons to 500–850 hPa (Fig. 2b). There is now a clear distinction between CO₂-only signals and signals that additionally incorporate aerosol effects. Annual $R(t)$ time series show overall positive trends for the SC-TP, SC-HC, S-TP, COMB1 and COMB2 signals, but little or no trend for the CO₂-only signals and a smaller positive trend for COMB3. The primary reason for this discrimination is the inter-hemispheric asymmetry common to the observations and the signals with combined CO₂ + aerosol forcing. This common asymmetry occurs also in DJF and JJA.

Although visibly apparent (Fig. 1), the benefit of incorporating stratospheric O₃ effects is not clearly shown in the $R(t)$ results. It becomes more obvious in an uncentred statistic, which retains the spatial means of the two fields being compared. This, however, has other limitations in the context of attributing observed changes to a specific causal factor^{11,37}.

Trend significance

Are the positive $R(t)$ trends in Fig. 2 unusually large relative to the trends we might expect in the absence of any anthropogenic forcing—that is, due to natural variability of the climate system? To address this issue, we use (internally generated) natural variability information from multi-century CGCM control experiments performed at the Hadley Centre (HC)¹⁰, the Max-Planck Institute for Meteorology in Hamburg (MPI)³⁸ and GFDL³⁹. All integrations were run with no changes in natural external or anthropogenic forcings.

A number of studies have attempted to assess how reliably these three CGCMs simulate real-world natural variability^{39–42}. On timescales of 10–30 years there is good agreement between the GFDL and HC spectra and the observed spectrum for global-mean annually averaged near-surface temperature^{39,43}. Comparisons of model and observed patterns of surface temperature variability show that similarities exist on timescales of 5–10 years, although there are differences on shorter timescales^{39,41,44}.

More rigorous validation of the model-estimated internal variability of vertical temperature changes is difficult because the radiosonde record is short and represents a convolution of natural variability and anthropogenic effects. For the purposes of this investigation we must therefore assume that the CGCMs used here provide credible estimates of the spectrum and patterns of

internal natural variability on timescales ranging from 10 to 25 years. Our use of noise information from multiple control runs provides an indication of the robustness of our significance estimates to uncertainties in the model-estimated noise.

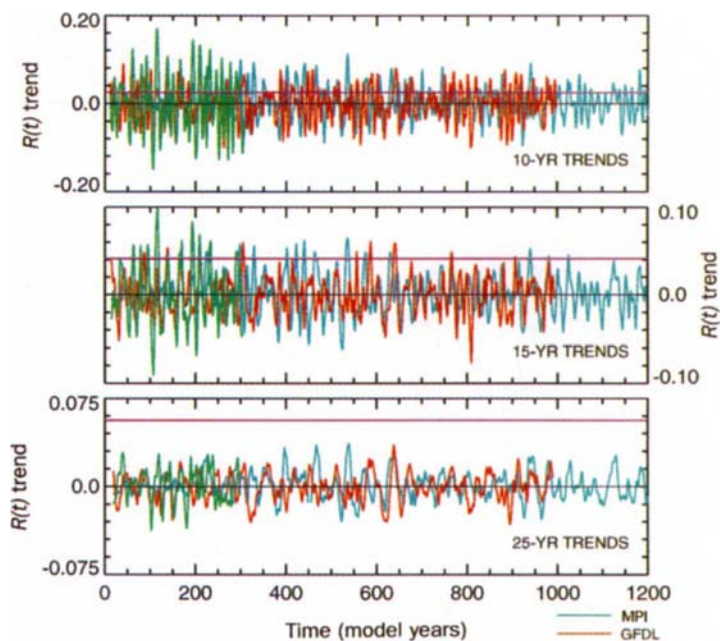
In the significance-testing procedure¹¹, each model signal pattern (with the exception of the O₃ and S-TP signals) is correlated with the sequences of vertical temperature-change patterns simulated in each of the three CGCM control runs. The resulting time series provide information on the behaviour of the $R(t)$ statistic in the absence of external forcing. By fitting 10-, 15-, and 25-year linear trends to overlapping chunks of these 'natural' $R(t)$ time series, we generate sampling distributions of unforced $R(t)$ trends. The 'signal' trends in $R(t)$ over the past 10 to 25 years (Fig. 2a, b) are then compared with these sampling distributions to determine whether changes in model–observed pattern similarity with time are statistically unusual (Fig. 3).

For model–observed comparisons over 50–850 hPa, the positive 25-year $R(t)$ trends for all signals considered (CO₂-only, CO₂ + aerosols, and the three sensitivity studies) differ significantly from unforced trends (Table 1a). This result shows that the observed pattern of change over 1963–87 is similar to model-predicted signals that incorporate CO₂ effects, whereas 'natural' patterns of change over 25-year periods are not⁴⁵.

On shorter timescales (10 and 15 years) signal-to-noise ratios decrease and significance results show a seasonal dependence, with fewer significant trends for both JJA and annually averaged data. There is also a dependence on the CGCM used to estimate noise properties, with higher noise levels (see Fig. 3) and fewer significant results for noise estimates obtained from the HC control run. The reverse applies for the MPI noise estimates. The 10- and 15-year results show why natural variability makes it difficult to evaluate the significance of short timescale trends, such as the lower tropospheric temperature changes estimated since 1979 from the satellite-based Microwave Sounding Unit^{46,47}.

Restricting pattern comparisons to 500–850 hPa yields a clear discrimination between the CO₂-only signals and signals that incorporate aerosol effects (Table 1b). In DJF and in the annually averaged data, the 25-year $R(t)$ signal trends are highly significant for the SC-TP, SC-HC, COMB1 and COMB2 signals, but not for COMB3 (except for annually averaged data relative to the MPI noise) or the CO₂-only signals. As noted above, this result arises

FIG. 3 The relationship between forced and unforced trends in the $R(t)$ pattern similarity statistic. 'Unforced' $R(t)$ time series were computed by correlating the fixed pattern of annually-averaged vertical temperature changes in the SC-TP experiment (Fig. 1c) with the time-varying temperature-change patterns from the 310-, 1,000- and 1,260-year HC, GFDL, and MPI control integrations (respectively). Model anomaly patterns were defined relative to the overall time-mean of each control run, and were filtered in the same way as the observations (see Fig. 2). The figure shows the result of fitting overlapping linear trends to 10-, 15- and 25-year segments of the unforced $R(t)$ time series, and then plotting the magnitude (at any point in time) of the linear trend in $R(t)$. The overlap between trends is by all but one year. This yields sampling distributions of all possible unforced $R(t)$ trends for the three selected timescales. The purple horizontal lines in each panel give the magnitude of the $R(t)$ trend for the comparison of the SC-TP signal with observations over 1978–87, 1973–87 and 1963–87 (see Fig. 2a). All trends are expressed as the change in $R(t)$ per year. The level of time-increasing similarity between the observed vertical temperature-change patterns and the SC signal over the last 25 years is highly unusual relative to the unforced 25-year $R(t)$ trends estimated from all three control integrations. In contrast, recent 10-year trends in $R(t)$ are not unusual occurrences. Note that the magnitude of the unforced $R(t)$ trends decreases as the length of $R(t)$ trends increases. The results shown here are for annually averaged data and for temperature-change patterns defined over 50–850 hPa. $R(t)$ trends are plotted on the central year of the trend.



mainly from the relatively lower Northern Hemisphere warming that is common to the observations and the combined forcing signals. Again, none of the shorter-timescale trends in $R(t)$ achieve significance at the 5% level or better.

Equilibrium and transient signal similarity

We have shown above that the TP equilibrium and HC transient signals yield very similar detection results when compared with the observed radiosonde data. This does not necessarily signify that the TP and HC signals are *themselves* correlated. Such similarity between equilibrium and transient signal patterns is implicitly assumed in any detection study that makes use of equilibrium signals⁴⁸. To test this assumption in a rigorous way would require signal estimates from both types of experiment performed with the same model. These were not available here. We gain some indication of the similarity between equilibrium and transient results by computing pattern correlations (see equation (1)) between the TP and HC signals over the 50–850 hPa and 500–850 hPa domains. For the full vertical domain, the key correlations for annually averaged data (and for HC signals that are decadal averages for the 1990s) are $R\{C-TP:C-HC\} = 0.87$; $R\{SC-TP:SC-HC\} = 0.85$. For the low- to mid-troposphere the key result is $R\{SC-TP:SC-HC\} = 0.61$. Note, however, that the CO₂-only signals in the TP and HC experiments are uncorrelated over 500–850 hPa ($R\{C-TP:C-HC\} = -0.05$).

A further assumption underlying our detection strategy is that the transient signal is relatively stable with time over the period relevant for a comparison with observations⁴⁸. We tested this by computing correlations between decadal averaged HC signal patterns for the modelled '1980s' and '1990s' $R\{C-HC80:C-HC90\} = 0.97$; $R\{SC-HC80:SC-HC90\} = 0.98$ (for 50–850 hPa) and $R\{SC-HC80:SC-HC90\} = 0.93$ (for 500–850 hPa).

These results illustrate that the signal components of most interest here—stratospheric cooling/tropospheric warming (in experiments that incorporate CO₂ effects) and hemispheric-scale temperature asymmetry (in experiments that include aerosol effects)—are similar in the TP and HC experiments and are relatively stable over the time period of the HC integrations relevant for comparison with the radiosonde data. We note, however, that these findings are not necessarily of general applicability to different models, climate variables and geographical domains.

Conclusions

Our results suggest that the similarities between observed and model-predicted changes in the zonal-mean vertical patterns of temperature change over 1963–87 are unlikely to have resulted from natural internally generated variability of the climate system. This conclusion holds for pattern comparisons over 50–850 hPa, which focus on the large-amplitude signal of stratospheric cooling and tropospheric warming, and for comparisons over 500–850 hPa, which emphasize hemispheric-scale temperature asymmetries in the lower atmosphere. Stratospheric cooling and tropospheric warming are common to the observations and all signals that incorporate CO₂ effects, whereas hemispheric asymmetry is an observed feature that is found only in model experiments that incorporate aerosol effects.

The main uncertainties in our work relate to:

(1) Estimates of the relative magnitudes and spatial and temporal evolution of the different forcings^{29,30}, including those human factors represented here (greenhouse gases, direct sulphate aerosol effects and stratospheric ozone), those not specifically represented here (such as indirect aerosol effects^{31,32,49}, other anthropogenic aerosols^{50–52}, tropospheric ozone and other non-CO₂ greenhouse gases⁵³), and purely natural forcings (for example, variations in solar output and volcanic aerosol loadings).

(2) The simulated response to forcing, including model-dependent factors such as global and regional sensitivity, which in turn depend on model parametrizations (for example, for clouds,

which directly affect global sensitivity⁵⁴, and for oceanic vertical mixing, which may affect interhemispheric asymmetry).

(3) The realism of CGCM-derived estimates of natural internal variability on decadal and longer timescales⁵⁵, and the neglect of the variability arising from natural external forcings.

(4) The existence of time-varying instrumental biases in the radiosonde data, and their incomplete spatial coverage^{7,8,56,57}.

Where possible, we have attempted to explore the sensitivity of our principal results to these uncertainties. For comparisons focusing on the pattern of stratospheric cooling and tropospheric warming, we have shown that the significant 25-year $R(t)$ trends are seasonally robust. They are insensitive to uncertainties in the signals examined here (as shown by our use of equilibrium and transient signals from models with differences in physical processes and in their treatment of direct aerosol effects), in the magnitude of the direct aerosol forcing (as shown by COMB3), and in the incorporation of stratospheric O₃ effects (as shown by COMB1 and COMB2). Finally, they are robust to uncertainties in the CGCM estimates of natural internal climate variability used here—to achieve non-significant results on 25-year timescales would require noise levels for 'unforced' $R(t)$ trends to be roughly twice as large as those estimated here.

The situation is more equivocal for model-observed comparisons focusing on hemispheric temperature-change asymmetries in the lower atmosphere. Here, too, the 25-year $R(t)$ temperature trends are generally robust to CGCM differences in natural internal variability and to signal uncertainties arising from model differences in physical processes and in the representation of direct aerosol effects. They are, however, sensitive to uncertainties in the magnitude of the aerosol forcing (as shown by COMB3). Although this points towards the need for caution in the interpretation of our results, we note that the observations are in better agreement with the larger aerosol forcing and hemispheric asymmetry in COMB1 and COMB2 than with the reduced forcing and asymmetry in COMB3. This may reflect our effective incorporation of some of the pattern characteristics of indirect aerosol forcing in the response to direct forcing²⁷. For this to be the case would require a hemispherically asymmetric indirect aerosol effect, and there is evidence to support this in recent modelling studies^{31,32} and satellite observations of interhemispheric differences in cloud liquid-water droplet size³³.

This investigation shows the clear need for modelling experiments with simultaneous changes in CO₂, O₃ and anthropogenic sulphate aerosols. Our use of linear superposition to assess the possible effects of ozone changes on a CO₂ + aerosol signal should be regarded as a sensitivity study only. The incorporation of stratospheric O₃ effects in this way improves the fit with observations by reducing the height of the transition from stratospheric cooling to tropospheric warming. The implication of this result and other recent work^{14,15} is that climate-change detection investigations that ignore possible ozone effects are likely to be searching for a sub-optimal signal (at least in terms of vertical temperature changes).

There is scope to reduce uncertainties in the observational data. We tested the robustness of our results to these uncertainties by comparing the radiosonde data with a reanalysis of operationally produced climate data. At present, this comparison is only possible after 1979. The extension of this reanalysis back to the early 1960s will allow better evaluation of the reliability of long-term trends in the radiosonde data. This will also be facilitated by other reanalysis efforts⁵⁸ and compilations of quality-controlled radiosonde data^{57,59}.

Although we have identified a component of the observational record that shows a statistically significant similarity with model predictions, we have not quantified the relative magnitude of natural and human-induced climate effects. This will require improved histories of radiative forcing due to natural and anthropogenic factors, and numerical experiments that better define an anthropogenic climate-change signal and the variability due to purely natural causes. □

Received 9 April; accepted 30 May 1996.

1. Madden, R. A. & Ramanathan, V. *Science* **209**, 763–768 (1980).
2. Epstein, E. S. *J. appl. Met.* **21**, 1172–1182 (1982).
3. Karoly, D. J. *Geophys. Res. Lett.* **14**, 1139–1141 (1987).
4. Karoly, D. J. *Geophys. Res. Lett.* **16**, 465–468 (1989).
5. Schlesinger, M. E. & Mitchell, J. F. B. *Rev. Geophys.* **25**, 760–798 (1987).
6. Mitchell, J. F. B., Manabe, S., Meleshko, V. & Tokioka, T. in *Climate Change: The IPCC Scientific Assessment* (eds Houghton, J. T., Jenkins, G. J. & Ephraums, J. J.) 131–172 (Cambridge Univ. Press, 1990).
7. Karoly, D. J. *et al. Clim. Dyn.* **10**, 97–105 (1994).
8. Oort, A. H. & Liu, H. J. *Clim.* **6**, 292–307 (1993).
9. Taylor, K. E. & Penner, J. E. *Nature* **369**, 734–737 (1994).
10. Mitchell, J. F. B., Johns, T. C., Gregory, J. M. & Tett, S. F. B. *Nature* **376**, 501–504 (1995).
11. Santer, B. D. *et al. Clim. Dyn.* **12**, 77–100 (1995).
12. Prather, M. *et al. in Climate Change 1994: Radiative Forcing of Climate Change and an Evaluation of the IPCC IS92 Emission Scenarios* (eds Houghton, J. T. *et al.*) 73–126 (Cambridge Univ. Press, 1994).
13. Hansen, J. E., Lacis, A., Ruedy, R., Sato, M. & Wilson, H. *Nat. geogr. Res. Expl.* **9**, 142–158 (1993).
14. Hansen, J. E. *et al. Clim. Change* **30**, 103–117 (1995).
15. Ramaswamy, V., Schwarzkopf, M. D. & Randel, W. J. *Nature* (submitted).
16. Walton, J. J., MacCracken, M. C. & Ghan, S. J. *J. geophys. Res.* **93**, 8339–8354 (1988).
17. Penner, J. E., Atherton, C. A. & Graedel, T. E. in *Global Atmospheric-Biospheric Chemistry* (ed. Prinn, R.) 223–248 (Plenum, New York, 1994).
18. Taylor, K. E. & Ghan, S. J. *J. Clim.* **5**, 906–919 (1992).
19. Charlson, R. J. & Wigley, T. M. L. *Scient. Am.* **270**, 48–57 (1994).
20. Johns, T. C. *et al. Clim. Dyn.* (submitted).
21. Houghton, J. T., Callander, B. A. & Varney, S. K. (eds) *Climate Change 1992: The Supplementary Report to the IPCC Scientific Assessment* (Cambridge Univ. Press, 1992).
22. Mitchell, J. F. B., Davis, R. A., Ingram, W. J. & Senior, C. A. *J. Clim.* **8**, 2364–2386 (1995).
23. Karl, T. R. *Nature* **371**, 380–381 (1994).
24. Christy, J. R. *Clim. Change* **31**, 455–474 (1995).
25. Kalnay, E. *et al. Bull. Am. met. Soc.* **77**, 437–471 (1996).
26. Mahlman, J. D., Pinto, J. P. & Umscheid, L. J. *J. atmos. Sci.* **51**, 489–508 (1994).
27. Penner, J. E., Wigley, T. M. L., Jaumann, P., Santer, B. D. & Taylor, K. E. in *Communicating About Climate: the Story of the Model Evaluation Consortium for Climate Assessment* (eds Howe, W. & Henderson-Sellers, A.) (Gordon & Breach, in the press).
28. Ramaswamy, V. & Bowen, M. M. *J. geophys. Res.* **99**, 18909–18921 (1994).
29. Shine, K. P., Fouquart, Y., Ramaswamy, V., Solomon, S. & Srinivasan, J. in *Climate Change 1994: Radiative Forcing of Climate Change and an Evaluation of the IPCC IS92 Emission Scenarios* (eds Houghton, J. T. *et al.*) 163–203 (Cambridge Univ. Press, 1994).
30. Shine, K. P., Fouquart, Y., Ramaswamy, V., Solomon, S. & Srinivasan, J. in *Climate Change 1995: The Science of Climate Change* (eds Houghton, J. T. *et al.*) 108–118 (Cambridge Univ. Press, 1996).
31. Jones, A., Roberts, D. L. & Slingo, A. *Nature* **370**, 450–453 (1994).
32. Erickson, D. J., Oglesby, R. J. & Marshall, S. *Geophys. Res. Lett.* **22**, 2017–2020 (1995).
33. Boucher, O. *J. Clim.* **8**, 1403–1409 (1995).
34. Chuang, C. C., Penner, J. E., Taylor, K. E., Grossmann, A. S. & Walton, J. J. *J. geophys. Res.* (submitted).
35. Angell, J. K. *J. Clim.* **1**, 1296–1313 (1988).
36. Barnett, T. P. & Schlesinger, M. E. *J. geophys. Res.* **92**, 14772–14780 (1987).
37. Wigley, T. M. L. & Barnett, T. P. in *Climate Change: The IPCC Scientific Assessment* (eds Houghton, J. T., Jenkins, G. J. & Ephraums, J. J.) 239–255 (Cambridge Univ. Press, 1990).
38. Hasselmann, K. *et al. Rep. No. 168* (Max-Planck-Institut für Meteorologie, Hamburg, 1995).
39. Stouffer, R. J., Manabe, S. & Vinnikov, K. Ya. *Nature* **367**, 634–636 (1994).
40. Tett, S., Johns, T. C. & Mitchell, J. F. B. *Clim. Dyn.* (submitted).
41. Manabe, S. & Stouffer, R. J. *J. Clim.* **9**, 376–393 (1996).
42. Hegerl, G. C. *et al. J. Clim.* (in press).
43. Santer, B. D., Wigley, T. M. L., Barnett, T. P. & Anyamba, E. in *Climate Change 1995: The Science of Climate Change* (eds Houghton, J. T. *et al.*) 407–443 (Cambridge Univ. Press, 1996).
44. Gates, W. L. *et al. in Climate Change 1995: The Science of Climate Change* (eds Houghton, J. T. *et al.*) 229–284 (Cambridge Univ. Press, 1996).
45. Vinnikov, K. Ya., Robock, A., Stouffer, R. J. & Manabe, S. *Geophys. Res. Lett.* (in the press).
46. Christy, J. R. & McNider, R. T. *Nature* **367**, 325 (1994).
47. Jones, P. D. *Geophys. Res. Lett.* **21**, 1149–1152 (1994).
48. Barnett, T. P., Schlesinger, M. E. & Jiang, X.-J. in *Greenhouse-Gas-Induced Climatic Change: A Critical Appraisal of Simulations and Observations* (ed. Schlesinger, M. E.) 537–558 (Elsevier, Amsterdam, 1991).
49. Wigley, T. M. L. *Nature* **339**, 365–367 (1989).
50. Penner, J. E., Dickinson, R. & O'Neill, C. *Science* **256**, 1432–1434 (1992).
51. Tegen, I. & Fung, I. *J. geophys. Res.* **100**, 18707–18726 (1995).
52. Penner, J. E. *et al. Bull. Am. met. Soc.* **75**, 375–400 (1994).
53. Wang, W.-C., Dudek, M. P., Liang, X.-Z. & Kiehl, J. T. *Nature* **350**, 573–577 (1991).
54. Mitchell, J. F. B., Senior, C. A. & Ingram, W. J. *Nature* **341**, 132–134 (1989).
55. Barnett, T. P., Santer, B. D., Jones, P. D., Bradley, R. S. & Briffa, K. R. *Holocene* (in the press).
56. Gaffen, D. J. *J. geophys. Res.* **99**, 3667–3676 (1994).
57. Parker, D. E. & Cox, D. I. *Int. J. Climatol.* **15**, 473–496 (1995).
58. Trenberth, K. E. *Clim. Change* **31**, 427–453 (1995).
59. Eskridge, R. E. *et al. Bull. Am. met. Soc.* **76**, 1759–1775 (1995).

ACKNOWLEDGEMENTS. We thank J. Boyle, M. Fiorino, D. Gaffen, L. Gates, T. Karl, J. Mahlman, S. Manabe, A. Robock, H. Rodhe and four reviewers for helpful suggestions; L. Corsetti (PCMDI) for providing programming assistance; J.-S. von Storch and G. Hegerl for supplying data from the MPI control integration; and D. Williams, R. Mobley and R. Drach for developing software used to produce the colour graphics. Work at Lawrence Livermore National Laboratory was performed under the auspices of the US Department of Energy, Environmental Sciences Division. Support for T.J., J.M. and S.T. was provided by the UK Department of the Environment. T.M.L.W. and P.D.J. were supported by the US Department of Energy, Environmental Sciences Division. T.M.L.W. was also supported under the NOAA Climate Change, Data and Detection Program. The MECCA Program supplied some of the computer time required for the TP integrations.

CORRESPONDENCE should be addressed to B.D.S. (e-mail: bsanter@rainbow.llnl.gov).

YOURS TO HAVE AND TO HOLD BUT NOT TO COPY

The publication you are reading is protected by copyright law. Photocopying copyright material without permission is no different from stealing a magazine from a newsagent, only it doesn't seem like theft.

If you take photocopies from books, magazines and periodicals at work your employer should be licensed with CLA.

Make sure you are protected by a photocopying licence.



The Copyright Licensing Agency Limited
90 Tottenham Court Road, London W1P 0LP
Telephone: 0171 436 5931 Fax: 0171 436 3986

Near-infrared jets in the Galactic microquasar GRS1915+105

Bruce J. Sams*, Andreas Eckart* & Rashid Sunyaev†

* Max-Planck-Institut für extraterrestrische Physik, Giessenbachstrasse, 85748 Garching, Germany

† Max-Planck-Institut für Astrophysik, Karl-Schwarzschild Strasse 1, 85748 Garching, Germany and Space Research Institute, Profsoyuznaya 84/32, Moscow, 117 810, Russia

QUASARS and active galactic nuclei often have radio jets emanating from them, parts of which sometimes appear to move faster than light¹. These jets are thought to arise from processes associated with a supermassive black hole, but it is difficult to study them in detail both because they lie at extreme distances and because observable changes in them occur only slowly. The Galactic 'microquasar' GRS1915+105 has attracted much attention because it also has superluminal jets²; it might be a low-power analogue of the 'central engine' of quasars, but so far the only evidence that it contains a black hole is its X-ray luminosity, which is greater than that allowed for a neutron star. Here we report observations of near-infrared jets from GRS1915+105; these jets emerge at the same position angle as the radio jets, but the luminosity associated with them is far greater. Using these data, we find a correlation between the length of a jet, its brightness temperature and the central source luminosity which applies across the entire range of black-hole-driven jets, thereby connecting microquasars, active galactic nuclei and quasars. Because the evolution of GRS1915+105 can be studied on relatively short timescales, we effectively have the ability to investigate evolutionary processes in quasars.

Since its recent discovery³, GRS1915+105 has been studied extensively in the radio and X-ray regimes, but because it lies 12.5 kpc distant in the Galactic plane behind $A_v \approx 20$ of dust extinction² optical-band imaging is impossible. We applied near-infrared speckle imaging and speckle polarimetry⁴ at $2.2 \mu\text{m}$ (K band, where $A_{2.2\mu\text{m}} = 0.112 A_v$) to obtain high-spatial-resolution (nearly diffraction limited) images of the GRS1915+105 counterpart⁵. We observed on 18–21 July 1995 using the Max Planck Institut für extraterrestrische Physik near-infrared speckle camera SHARP (ref. 6) at the New Technology Telescope of the European Southern Observatory. Data were taken in speckle, speckle polarimetric and direct modes. Our 12.8-arcsec field of view 0.05 arcsec per pixel was rotated -20° in position angle (PA), in order to cover both GRS1915+105 and two isolated stars ~ 10 arcsec to the northeast, which we call stars 1 and 2 (see Fig. 1 in this Letter and Fig. 3 in ref. 5). On 18 July the atmospheric seeing was 0.6 arcsec, which allowed speckle imaging with frames of 1 second exposure time. These frames contain high spatial frequencies which we recovered with a shift-and-add algorithm⁷. All observations used a K ($2.2 \mu\text{m}$) filter. Polarization intensities were obtained by rotating a polarizing element in increments of 30° (ref. 4). Our photometric reference was the Galactic Centre source IRS7, which has a K magnitude of 6.70 (ref. 8). Our absolute flux calibration error is $\lesssim 10\%$, or about 0.1 mag.

The final K-band map of the entire field (Fig. 1) shows the isolated stars and GRS1915+105, which appears extended to the southwest. To verify that the extension is not due to isoplanatism effects, we reduced the data in two ways: first, using both of the isolated stars as shift references, and second using the stars and GRS1915+105 as shift references. Both methods produced nearly identical raw images with limiting spatial resolution of 0.25 arcsec (diffraction limit, 0.15 arcsec).

Removing a point source from the image of GRS1915+105

reveals the jet structures more clearly (Fig. 1 inset). Their morphology is strikingly similar to that observed in a radio outburst². The entire source has $K = 12.5$ mag, which makes it 0.5 mag brighter than at any previously measured time^{5,9,10}. The southwest jet has a total near-infrared magnitude of $K = 13.9$, slightly brighter than the total source magnitude at its weakest measured value of $K = 14.3$ mag on 4 June 1993 (ref. 5), which shows that the lifetime of the near-infrared jets is less than 2 years. We do not detect a clear maximum corresponding to the northeast jet. We note, however, that the residuals to the southwest are 3.3 times brighter than those to the northeast, in agreement with the flux ratio of $\approx 4:1$ measured for the radio jets². For the likely jet ejection speed and geometry (ref. 2), the relativistic Doppler effect leads to redshifted emission from both jets. The redshifts are ~ 0.75 and ~ 2.36 for the approaching and receding jets, respectively. There are no reports of a pair of redshifted $H\alpha$ ($\lambda = 0.65 \mu\text{m}$) emission lines from GRS1915+105 as seen in SS433 (ref. 11), but even if there were such emission, it would not affect the observed southwest jet brightness, because $H\alpha$ emission is not redshifted into the K passband. However, $H\alpha$ might contribute to the northeast jet.

We determine the position angle of the southwest jet to be $140 \pm 10^\circ$, consistent with radio maps². The PA of the northeast extension is difficult to estimate, but the overall source extension is broadly consistent with the radio PA of $\sim 330^\circ$. The separation of the central source and the southwest jet component is 0.3 arcsec, which corresponds to 6×10^{16} cm, well above the dimensions of any bright X-ray binary system observed until now.

Polarimetry was accomplished using the two isolated stars (polarization $\lesssim 2\%$; refs. 5,12) in our field as differential

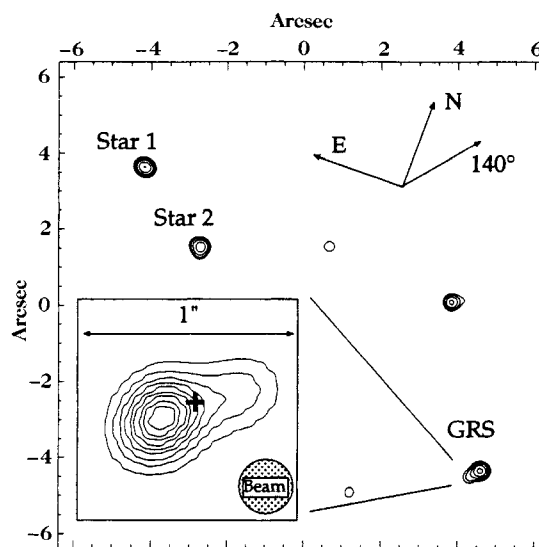


FIG. 1 K-band map of the 12.8×12.8 arcsec field containing GRS1915+105 (bottom right); its extension to the southwest is clearly visible. Two shift-and-add reference sources are at the upper left. This figure was created by the following technique (described in detail in ref. 23): after shift-and-add processing, a point-spread function was synthesized by coadding the two reference stars, and the image was deconvolved for 5,000 iterations using the Lucy–Richardson algorithm. The resulting map consists solely of delta functions. The delta function map was then convolved with a gaussian of resolution 0.25 arcsec to create the final image. Removing the delta functions associated with the central point source of GRS1915+105 (marked with a cross in the inset), reveals residual with a clear extension to the southwest. This is the southwest jet. No such residual appears when the same procedure is applied to the reference stars. The inset covers a one-arcsec region around the source. Contours in the inset are at 10, 20, 30 ... 100% of maximum. The jets are unresolved in our beam.

references to calibrate out any instrumental polarization and/or flux variations due to atmospheric effects. We find a total source polarization of $2.7 \pm 0.6\%$ oriented at $PA\ 64 \pm 5^\circ$, consistent with the expected polarization due to magnetic orientation of interstellar dust grains towards the Galactic Centre⁴. This indicates that the total near-infrared emission from GRS1915+105 is not intrinsically highly polarized. The jets alone could be polarized by up to 30% without contradicting our measurements at the 3σ level.

During our observations we detected no statistically significant flux variations on night-to-night or seconds-to-minutes timescales. This constancy contrasts strongly with 4 June 1993, when GRS1915+105 brightened in the K band 1.0 mag over a 24-hour period⁵. Our observations were made during an intense X-ray outburst when the X-ray flux increased by at least a factor of ~ 10 from its quiescent value of $F_X \lesssim 100$ mCrab on 16 June 1995 to $F_X \approx 500$ mCrab on 17–21 July 1995 (refs 13,14). There were no published X-ray observations between June and July.

Assuming a spectral index of $\alpha = 0.5$, the jet component we observe in the near-infrared with $K = 13.9$ mag through extinction of $A_V \approx 20$ should have a 3.5 cm flux of 770 mJy. This is 1.5 times higher than previous 3.5-cm measurements in March 1994 when the superluminal motion was detected², and where the young jets were observed to have $\alpha = 0.48$ (ref. 2). There is no published account of the radio flux on the date of our observations. For extragalactic sources the power-law spectrum is often self-absorbed at low frequencies: radio jets at the 20 mJy level would imply a power law index of $\alpha = 0$.

Two alternatives to optically thin synchrotron emission are dust in the plane of the binary system heated by the central object, and light echoes. The echoes could be reflections of very compact, bright and self-absorbed jets off surrounding dust and ionized gas. Light echoes could in principle be distinguished by the orientation and intensity of their polarization¹⁵. If there were simultaneous observations which showed no radio jets, they might suggest such an alternative model.

Systems dominated by a black hole of mass M have a single fundamental parameter with the dimension of length, namely the gravitational radius, $R_g = 2GM/c^2$. All lengths may thus be expressed in nondimensional units of R_g . Furthermore, the black-hole mass (m) and accretion rate ($\dot{m}$) can be expressed in non-dimensional units of the solar mass and Eddington luminosity ($L_{\text{Edd}} = 1.4 \times 10^{38} M/M_\odot \text{ erg s}^{-1}$) as $m = M/M_\odot$ and $\dot{m} = L/L_{\text{Edd}}$ respectively. The black-hole/accretion-disk bolometric luminosity is then $L_{\text{Bol}} = L_{\text{Edd}} \dot{m} \propto m\dot{m}$. In these units, we now compare the observed physical characteristics of microquasar jets with those of jets from active galactic nuclei and quasars (Fig. 2). Our comparison is phenomenological, not theoretical (see, for example, refs 16,17). Our sample consists of all well known massive black hole powered sources with both optical and radio jets and with published central source ROSAT or Einstein Observatory X-ray fluxes. To this we add GROJ1655–40 which is the only other known microquasar, but for which no optical jet has yet been detected. X-ray fluxes for the microquasars are from broad-band measurements³⁶. This sample is small but well defined, and is dominated by bright jet sources. We find two clear trends.

First, the length of the jets, l_{jet} , obeys $l_{\text{jet}} \propto L_{\text{Bol}}^{0.96} \propto m\dot{m}$. We take l_{jet} to be the length of the compact core, which is the length at which the jet surface brightness drops either below the detection limit, or to a level typical of any broad, diffuse emission. Obviously there are many factors (such as magnetic fields and ambient gas density) that can affect l_{jet} , but the correlation is clear for both optical and radio jets, and allows a simple prediction of the jet length: $l_{\text{jet}} \approx 3 \times 10^{10} R_g \dot{m}$. Second, the brightness temperature of the jets averaged over their surface areas obeys $T_b = A(m\dot{m})^{-0.8} \propto L_{\text{Bol}}^{-0.8}$, where the constant A is 3×10^7 K for $\lambda = 2$ cm, and 2×10^{-6} K for B band. Although well resolved radio maps of the jets in 3C273 and M87 show that a significant fraction of the flux comes from bright knots, to make a fair

comparison with the unresolved jets in GRS1915+105 and with marginally resolved jets in other sources, we take an average brightness value. We note a simple argument to explain the behaviour of T_b . The synchrotron volume emissivity of relativistic electrons with a power-law energy distribution given by $N(E) = NE^{-s}$ is $j_\nu \propto NB^{(s+1)/2} \nu^{-(s-1)/2}$ where E is the electron energy, ν is the emission frequency, and $s = 2\alpha - 1$ relates the spectral index to the slope of the electron spectrum¹⁸. Under conditions of equipartition, the energy density of cosmic rays is $\sim B^2/8\pi$, so $N \propto B^2$, and $j_\nu \propto B^{2s+3} \nu^{-s}$. For optically thin radiation, the jet surface brightness is $I_\nu = j_\nu d$ where d is the jet diameter. Assuming $d \propto R_g \dot{m}$ (as we found for l_{jet}) then $d \propto m\dot{m}$. The brightness temperature thus obeys $T_b \propto dB^{2s+3} \propto m\dot{m} B^{2s+3}$. From Fig. 2 we find $T_b \propto (m\dot{m})^{-0.8}$, so the jet magnetic field is given by $B \propto (m\dot{m})^{-1.8/(2s+3)}$. For $\alpha = 0.5$ we have $B \propto (m\dot{m})^{-1/2}$. Let us remark that the standard theory of accretion disks¹⁹ predicts that the maximum possible magnetic field at a given R/R_g in a radiation-dominated disk is $B \propto m^{-1/2}$. Consequently, the microquasar jets have B fields $\sim 10^4$ stronger than those in quasars, and their electrons need only be 10^{-2} as energetic to produce photons of the same observed frequency. Regardless of B , the electrons creating the near-infrared jet emission from GRS1915+105 are $\sqrt{\nu_K/\nu_{\text{radio}}} \sim 100$ times more energetic than those radiating the radio emission (where ν_K and ν_{radio} are the frequencies of K-band and radio radiation); the lifetimes are ~ 100 times shorter.

In any black-hole accretion model, the characteristic timescale

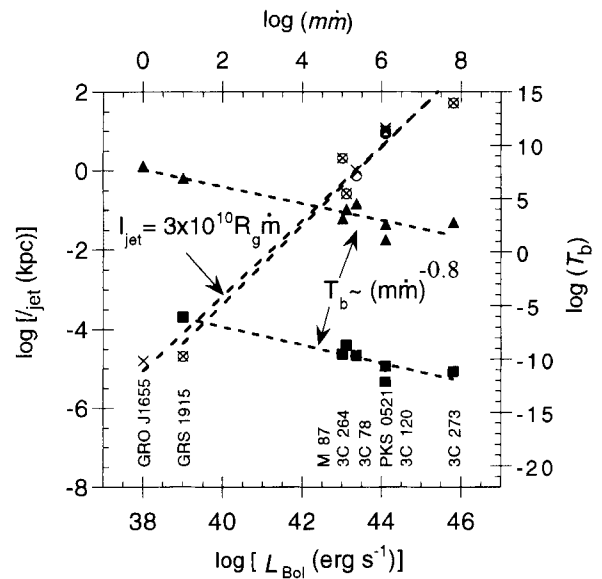


FIG. 2 A comparison between l_{jet} , T_b , L_{Bol} and $m\dot{m}$ for GRS1915+105, GROJ1655–40, five active galactic nuclei and a quasar. The L_{Bol} is V-band plus ROSAT or Einstein luminosity. The circles mark optical lengths, the crosses mark radio lengths, triangles mark optical brightness temperatures, and squares mark radio brightness temperatures. For GRS1915+105, where the optical jets are not clearly resolved, T_b is a lower limit. The radio and optical brightness temperatures have all been corrected (assuming power-law radiation with slope $\alpha = 0.5$) to 2-cm and B-band respectively. M87 is known to have a central black hole mass of $\sim 2 \times 10^9 M_\odot$ (ref. 37) which implies that its present X-ray emission is less than 10^{-4} of the Eddington limit. This supports the correlation with central source luminosity rather than with black hole mass. The L_{Bol} for 3C120 and PKS0521 are nearly identical, as are the jet lengths: T_b in the 3C120 jets is lower. Data for GRS1915+105 from this work and ref. 2; for GROJ1655–40 from refs 24 and 21; for 3C120 from ref. 25; for 3C78 from refs 26 and 27; for 3C264 from refs 28, 29 and 30; for PKS0521–36 from refs 31 and 32; for M87 from refs 33 and 32; for 3C273 from refs 34 and 35; some data also from ref. 28. Relations between the four plotted parameters have been derived from the slope of best-fit lines (shown dashed): $T_{b,r} \propto L_{\text{Bol}}^{-0.76}$, $T_{b,o} \propto L_{\text{Bol}}^{-0.84}$, $l_{\text{jet,r}} \propto L_{\text{Bol}}^{0.93}$, $l_{\text{jet,o}} \propto L_{\text{Bol}}^{0.99}$. Here $T_{b,r}$ and $T_{b,o}$ are the radio and optical brightness temperatures (upper and lower curves); $l_{\text{jet,r}}$ and $l_{\text{jet,o}}$ are the observed jet lengths in radio and optical (upper and lower curves).

of most instabilities in the accretion disk²⁰ is proportional to M , so the 1-day variability of the GRS1915+105 jets corresponds to $\sim 300,000$ years in the case of massive quasars such as 3C273. The trends presented here demonstrate that it is possible to compare the two types of systems directly. The proximity of GRS1915+105, together with the presence of at least two^{21,22} such sources in the Galaxy and the advent of high-resolution ground- and space-based imaging set the stage for studying the evolution of superluminal jets in much less than a human lifetime. □

Received 24 January; accepted 6 May 1996.

1. Begelman, M. C., Blandford, R. D. & Rees, M. J. *Rev. mod. Phys.* **56**, 255–351 (1984).
2. Mirabel, I. F. & Rodríguez, L. F. *Nature* **371**, 46–48 (1994).
3. Castro-Tirado, A. J., Brandt, S. & Lund, N. *IAU Circ. No. 5590* (1992).
4. Eckart, A., Genzel, R., Hofmann, R., Sams, B. J. & Tacconi-Garman, L. E. *Astrophys. J.* **445**, L23–L26 (1995).
5. Mirabel, I. F. et al. *Astr. Astrophys.* **335**, L17–L20 (1994).
6. Hoffmann, R. et al. in *Prog. in Telescope and Instrumentation Tech.* (ed. Ulrich, M. H.) 617–620 (ESO Conf. Ser. 42, ESO, Garching, 1993).
7. Christou, J. C. *Expl. Astr.* **21**, 27–32 (1991).
8. DePoy, D. L. & Sharp, N. A. *Astr. J.* **101**, 1324–1328 (1991).
9. Eikenberry, S. S. & Fazio, G. G. *IAU Circ. No. 6267* (1995).
10. Castro-Tirado, A. J. et al. *IAU Circ. No. 5590* (1992).
11. Margon, B. A. *Rev. Astr. Astrophys.* **22**, 507–536 (1984).

12. Axon, D. J. & Ellis, R. S. *Mon. Not. R. astr. Soc.* **177**, 499–511 (1976).
13. Sazonov, S. & Sunyaev, R. *IAU Circ. No. 6196* (1995).
14. Sazonov, S. & Sunyaev, R. *IAU Circ. No. 6209* (1995).
15. Gilfanov, M. R., Syunyaev, R. A. & Churazov, E. M. *Sov. Astr. Lett.* **13**, 233–239 (1988).
16. Falcke, H. & Biermann, P. L. *Astr. Astrophys.* **308**, 321–327 (1996).
17. Levinson, A. & Blandford, R. *Astrophys. J.* **456**, L29–L32 (1996).
18. Longair, M. S. *High Energy Astrophysics* (Cambridge Univ. Press, 1986).
19. Shakura, N. I. & Sunyaev, R. A. *Astr. Astrophys.* **24**, 337–355 (1973).
20. Shakura, N. I. & Sunyaev, R. A. *Mon. Not. R. astr. Soc.* **175**, 613–632 (1976).
21. Hjellming, R. M. & Rupen, M. P. *Nature* **375**, 464–468 (1995).
22. Bailyn, C. D., Orosz, J. A., McClintock, J. E. & Remillard, R. A. *Nature* **378**, 157–159 (1995).
23. Eckart, A. et al. in *The Nuclei of Normal Galaxies* (eds Genzel, R. & Harris, A.) 305–315 (Nato ASI Ser. 445, Kluwer Academic, Dordrecht, 1994).
24. Harmon, B. A. et al. *Nature* **374**, 703–706 (1995).
25. Hjorth, J., Vertergaard, M., Sørensen, A. N. & Grundahl, F. *Astrophys. J.* **452**, L17–L20 (1995).
26. Sparks, W. B. et al. *Astrophys. J.* **450**, L55–L58 (1995).
27. Unger, S. W., Booler, R. V. & Pedlar, A. *Mon. Not. R. astr. Soc.* **207**, 679–684 (1984).
28. Crane, P. et al. *Astrophys. J.* **402**, L37–L40 (1993).
29. Elvis, M., Schreier, E. J., Tonry, J., Davis, M. & Huchra, J. P. *Astrophys. J.* **246**, 20–27 (1981).
30. Gavazzi, G., Perola, G. C. & Jaffe, W. *Astr. Astrophys.* **103**, 35–43 (1981).
31. Keel, W. C. *Astrophys. J.* **302**, 296–305 (1986).
32. Keel, W. C. *Astrophys. J.* **329**, 532–550 (1988).
33. Biretta, J. A., Stern, C. P. & Harris, D. E. *Astr. J.* **101**, 1632–1646 (1991).
34. Biretta, J. A., Owen, F. N. & Comwell, T. J. *Astrophys. J.* **342**, 128–134 (1989).
35. Bahcall, J. N., Kirhakos, S., Schneider, D. P. & Unwin, S. C. *Astrophys. J.* **452**, L91–L93 (1995).
36. Sazonov, S. & Sunyaev, R. *IAU Circ. No. 6209* (1995).
37. Ford, H. C. et al. *Astrophys. J.* **435**, L27–L30 (1994).

ACKNOWLEDGEMENTS. We thank R. Genzel for his support and helpful critical discussions.

A possible high-temperature origin for the carbonates in the martian meteorite ALH84001

Ralph P. Harvey* & Harry Y. McSween Jr†

* Department of Geological Sciences, Case Western Reserve University, Cleveland, Ohio 44106-7216, USA

† Department of Geological Sciences, University of Tennessee, Knoxville, Tennessee 37996-1410, USA

THE meteorite Allan Hills (ALH) 84001, commonly accepted to be of martian origin, is unique among known martian meteorites in containing abundant, zoned, pre-terrestrial carbonate minerals^{1–9}. Previous studies of the oxygen isotope compositions of these minerals^{5,6,8} have suggested that they precipitated from a low-temperature (0–80 °C) aqueous fluid in the martian crust—perhaps in a near-surface hydrothermal system. Here we report analyses of the major-element compositions of the carbonates, which provide an independent constraint on the composition and temperature of the fluid from which they formed. We argue that the most likely explanation for the observed compositions, and for the absence of co-existing hydrous minerals, is that the carbonates were formed by reactions between hot (>650 °C), CO₂-rich fluids and the ultramafic host rock during an impact event. Impact processes on the martian surface can produce both the hot, CO₂-rich fluid (by volatilization of surface carbonates or other CO₂ sources) and—by brecciation—the conduits through which it flowed. Impact metasomatism is also consistent with the observed oxygen isotope disequilibrium, sequence of mineral formation, and carbonate mineral zoning, reflecting carbonate formation during rapid cooling from high temperatures rather than prolonged exposure to low-temperature fluids^{6,8}.

The martian meteorite ALH84001 is composed predominantly of centimetre-sized equant orthopyroxene grains (Wo₃En₇₀Fs₂₇) arranged in a contact lattice, with interstitial areas containing maskelynite, augite, chromite, pyrite and apatite^{3–5}. Crushed regions a few millimetres wide cut mostly through interstitial areas and contain a mixture of cumulus and interstitial phases as well as rarer phases such as olivine (slightly more Fe-rich than the orthopyroxene), sulphides and the Mg–Fe–Ca carbonates, which include calcite (CaCO₃), dolomite–ankerite (Ca (Mg,Fe) (CO₃)₂), and magnesite–siderite ((Mg,Fe)(CO₃)). Unambiguously extraterrestrial hydrous phases have been sought by many authors and are apparently absent^{3–5,9}.

The carbonates occur as fine-grained, spherical structures exhibiting radial Mg–Fe–Ca zoning (Fig. 1). The zoning pattern is identical for rosettes throughout the meteorite^{3–5,7} (Fig. 1): a general core–rim trend from Ca- and Fe-rich to more Mg-rich compositions is surrounded by a sharper micrometre-scale Fe-rich band, and is in turn mantled by a Mg-rich rim. Occasionally dolomite-calcite cores, a diffuse inner band of Ca,Fe-enrichment, or a second rim of Fe-rich carbonate can be seen when the plane of the thin-section intersects the rosette through its centre.

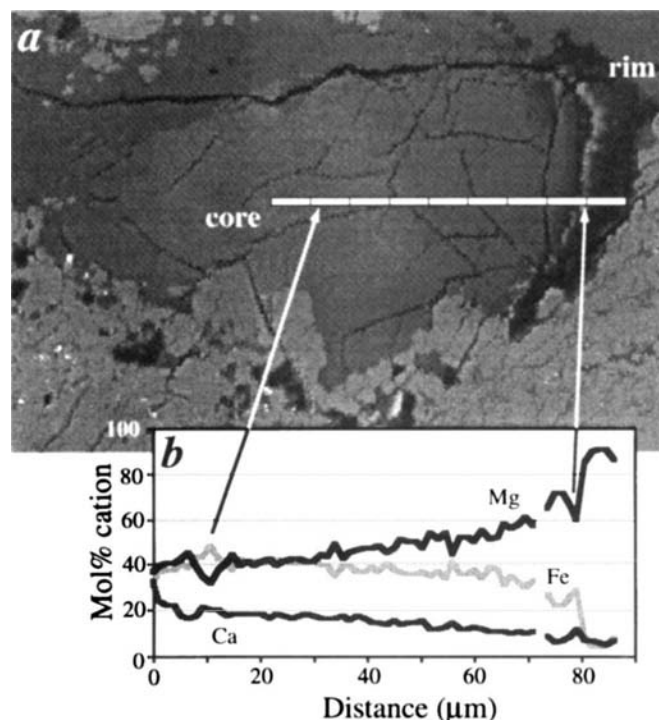


FIG. 1 Backscattered-electron image of a typical ALH84001 carbonate rosette (a) with zoning profiles (b). The largest volume consists of micrometre-sized magnesite–siderite grains exhibiting increasing Fe content with increasing radial distance. At several points on a typical transect this general trend is interrupted by bands of increased Fe or Ca content. A diffuse band of minor Fe increase (5–10%) typically occurs within $\sim 20 \mu\text{m}$ of the apparent centre, and a much sharper band of Fe-rich carbonate, only a few micrometres wide, occurs near the outer regions of the rosette. This outer, Fe-rich band is in turn mantled by an extensive rim of Mg-rich carbonates exhibiting little compositional variation.

The texture of the rosettes, their identical zoning and preferential siting in crushed zones argue for serial crystallization from a fluid circulating in interconnected pore spaces of this meteorite during the entire growth period. Although commonly found in proximity to (or surrounded by) maskelynite, carbonate rosettes are also found where maskelynite is absent. These rosettes also nucleate on orthopyroxene, apatite and other phases, suggesting kinetic control of growth as is found in many depositional systems. Massive, Mg-rich carbonate appears as a background in some areas where the margins of several rosettes have grown together (Fig. 2). Such areas sometimes contain silica at the juncture where several rosettes coalesce.

The major-element compositions of ALH84001 carbonates help constrain the composition and temperature of the fluid from which they formed. The substitution of Ca^{2+} for Mg^{2+} in the magnesite–siderite structure is rare below 300 °C, and remains limited even at 1,000 °C; as a result, significant Ca in magnesite–siderite carbonates without evidence of exsolution suggests formation at high temperature^{3,10}. In addition, the absence of solid solution between dolomite–magnesite and between calcite–siderite produces a unique three-carbonate assemblage at temperatures above 500 °C, with calcite and dolomite–ankerite forming a single continuous stability field for Fe-rich compositions and magnesite–siderite showing a wide solid-solution¹⁰. The compositions of more than 6,000 small (1–2 μm) carbonate regions, when superimposed on a CaCO_3 – MgCO_3 – FeCO_3 ternary (Fig. 3, left) show linear trends (due to pixels overlapping grain boundaries) between high Fe–Mg calcite, intermediate Mg–Fe dolomite, and a high Fe-member of the magnesite–siderite trend, consistent with a high-temperature carbonate assemblage. Fully quantitative analyses confirm the presence of distinct calcite, dolomite and magnesite–siderite phases (circles on Fig. 3, right). Calcite–dolomite thermometry¹⁰ for these analyses suggest formation at 680 °C; however, under metastable conditions this calculated temperature serves only as a maximum. In addition, the magnesite–siderite trend follows the high-temperature isotherm. This is commonly seen in terrestrial Mg–Fe carbonates where slight shifts in pH or Eh (redox state) dramatically shift the Mg/Fe ratio of depositing carbonates with little or no temperature change¹¹. The significance of the gentle curve shown by this trend is not known, and experimental data for carbonates of these compositions are nonexistent¹⁰.

Phase relations in the system CaO – MgO – SiO_2 – CO_2 – H_2O determined from both experimental and field studies show that a cooling, high-temperature fluid (>650 °C) with high values of X_{CO_2} (>0.85, where X_{CO_2} is the mole fraction of CO_2) will react with ultramafic minerals to produce calcite, dolomite and magnesite without producing significant volumes of hydrous minerals^{12–16}. Substitution of Fe for Mg in the mafic silicates does not appreciably change the phase relations¹³. The textural relationship of olivine, pyroxene and maskelynite in ALH84001^{4,5} suggests that these minerals were cation sources during carbonate production; however, because the expected products of feldspar–fluid reactions (for example, chlorite, spinel and anthophyllite) are missing, Ca may have other sources, such as apatite or augite, or the original fluid itself.

Impact-driven metasomatism, long hypothesized for Mars and known from Earth^{17,18} can mobilize volatiles and create conduits for their flow. ALH84001 apparently dates from a period of massive bombardment on Mars and has clearly experienced several severe shock events, including at least one major event preceding carbonate formation^{4,5,19}. Studies of terrestrial impact craters have shown that mobilized CO_2 can produce significant amounts of secondary carbonate during infiltration of impact-produced breccias¹⁸. Anhydrous CO_2 exists, in the form of surficial frosts, polar ice caps, permafrost and possible extensive deposits of carbonate within the martian soil²⁰. Impact-driven production of the ALH84001 carbonates is also consistent with previous studies of their C and O isotopic compositions^{6,8}. Enrichments in ^{13}C (+15 to +55‰) compared to terrestrial carbonates

suggest derivation at temperatures above 200 °C from a source in equilibrium with the martian atmosphere (measured at +27‰) and imply a surficial rather than deep-seated source⁶. Measured $\delta^{18}\text{O}$ values for ALH84001 carbonates are variable (+13 to +22‰) and not in equilibrium with silicate phases (+4.6‰)⁶. Although this discrepancy has been cited as evidence for a low-temperature (<100 °C) origin⁶, carbonates formed in terrestrial impacts show similar disequilibrium even when temperatures in excess of 1,200 °C are indicated by the presence of mafic glasses¹⁸. Alternatively the disequilibrium of oxygen isotopes in ALH84001 may represent exchange with a terrestrial reservoir, as also suggested by the possible presence of terrestrial minerals^{8,9}.

An impact-produced pulse of hot CO_2 -rich fluid would be expected to cool rapidly, allowing preservation of the observed fine-scale banding of ALH84001 carbonates that would diffuse away under prolonged high temperatures or exposure to an ambient fluid⁵. Rapid cooling also suggests that the products of lower-temperature reactions should be present in sequence at the exterior of carbonate rosettes. The silica found between the exteriors of coalescing carbonate rosettes⁴ supports this scenario, as it is an expected product of reactions between CO_2 and ultramafic rock at temperatures roughly 100 °C lower than those producing Mg–Fe carbonates^{12,13}. Reducing conditions at intermediate temperatures (probably <300 °C) have been invoked to explain micrometre-sized ZnS and Fe sulphates within the Fe-rich band at the rim of rosettes⁹. This band also encloses crystal faces of the interior magnesite, suggesting a possible hiatus in Mg–Fe carbonate growth such as might be expected in a short-lived stochastic system^{3,5}. The absence of hydrous minerals in ALH84001 makes it difficult to predict the products of reactions at still lower temperatures (<100 °C); recent models of possible weathering reactions for Mars have considered only hydrous conditions, and predict an assemblage rich in Fe hydroxides and clay minerals, in addition to carbonates and silica^{21,22}.

The concordance between ALH84001 carbonate compositions and the expected products of reactions between an initially high-temperature, CO_2 -rich (and H_2O -poor) fluid and an ultramafic host rock, and the lack of a corresponding volume of

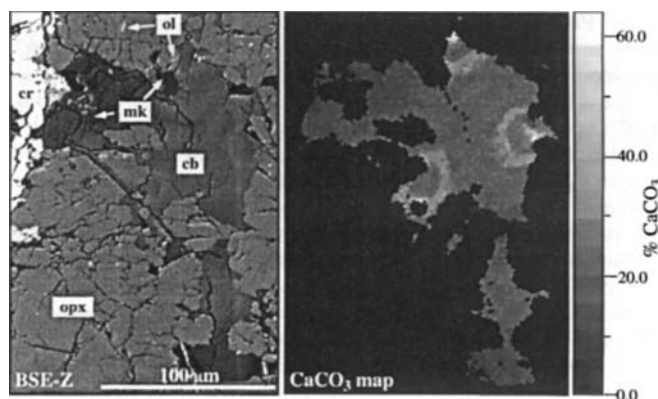
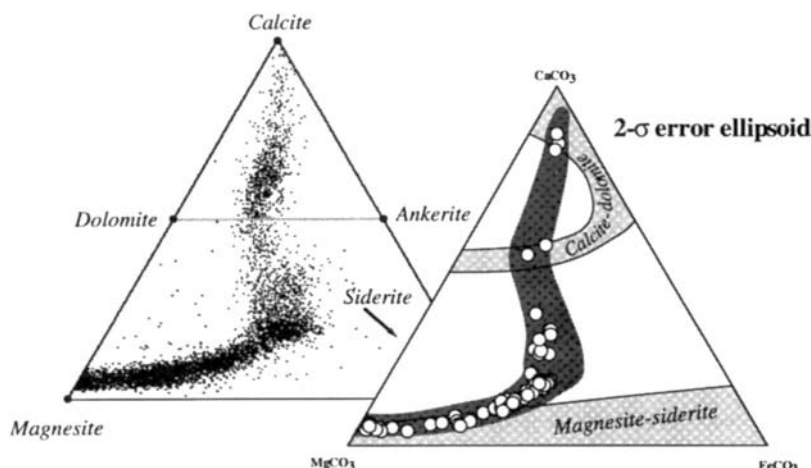


FIG. 2 Backscattered-electron image (left) and CaCO_3 image (right) of a representative carbonate area showing multiple growth centres. Phases surrounding carbonate (cb) include orthopyroxene (opx), chromite (cr), maskelynite (mk) and olivine (ol). Eight carbonate-bearing regions from two thin sections of ALH84001 were mapped at 1–3 μm resolution for 13 major elements using a Cameca SX-50 electron microprobe with Link automated imaging system. Each pixel in a 128×128 pixel grid was analysed for 500 mS at 20 nA and 10 keV. Raw X-ray count data for specific elements were gathered within specific energy dispersive windows, as well as background. These uncorrected elemental maps were then used to mask out non-carbonate phases (for example, to eliminate silicates, all pixels showing Si above background were masked). A partially corrected composition for each pixel was then produced by direct comparison to ZAF-corrected wavelength-dispersive analyses of identical sites for a few dozen selected pixels on each image. 2σ analytical sensitivity ranged from 2.5% for Mg to 1.4% for Fe in this procedure.

FIG. 3 Carbonate compositions in ALH84001. Left, 6,736 semi-quantitative analyses form two individual carbonate regions in ALH84001. Note that pixel size and carbonate grain size are nearly identical, so that many individual analyses represent compositions on the border of two grains. Right, 700 °C ternary phase diagram¹¹. Quantitative analyses are shown as white circles, dark-shaded area shows compositions seen on the left.



hydrous phases, suggest that impact-driven metasomatism may be an important source of carbonate mineral production on Mars. Martian carbonates, rather than being deposited by an actively circulating, long-lived water-rich system, may instead have been produced during isolated, short-term cycling of CO₂ within the ancient, impact-fractured crust. Over time this process may have globally redistributed carbonates throughout the shallow crust of Mars, possibly disguising carbonate deposits produced during collapse of an early CO₂-rich atmosphere. □

Received 5 October 1995; accepted 9 May 1996.

1. McSweeney, H. Y. Jr *Meteoritics* **29**, 757–779 (1994).
2. Score, R. & Lindstrom, M. *Ant. Meteorite Newsl.* **18**(2), 20 (1995).
3. Mittlefehdt, D. W. *Meteoritics* **29**, 214–221 (1994).
4. Harvey, R. P. & McSweeney, H. Y. Jr *Meteoritics* **29**, 472 (1994).
5. Treiman, A. H. *Meteoritics* **30**, 294–302 (1995).
6. Romanek, C. S. et al. *Nature* **372**, 655–657 (1994).
7. Romanek, C. S., Thomas, K. L., Gibson, E. K. Jr, McKay, D. S. & Socki, R. A. *Meteoritics* **30**, 567–568 (1995).

8. Jull, A., Eastoe, C. J., Xue, S. & Herzog, G. F. *Meteoritics* **30**, 311–318 (1995).
9. Wentworth, S. J. & Gooding, J. L. *Proc. Lunar planet. Sci. Conf.* **XXVI**, 1489–1490 (1995).
10. Anovitz, L. M. & Essene, E. J. *J. Petrol.* **28**, 389–414 (1987).
11. Woods, T. L. & Garrels, R. M. *Geochim. cosmochim. Acta* **56**, 3031–3043 (1992).
12. Trommsdorff, V. & Connolly, J. A. D. *Contr. Miner. Petrol.* **104**, 1–7 (1990).
13. Johannes, W. *Am. J. Sci.* **267**, 1083–1104 (1969).
14. Schreyer, W., Ohnmacht, W. & Mannchen, T. *Lithos* **5**, 345–364 (1972).
15. Cribb, S. J. *Norsk Geol. Tidsskrift* **62**, 161–168 (1982).
16. Ohnmacht, W. *J. Petrol.* **15**, 303–323 (1974).
17. James, P. B., Kieffer, H. H. & Paige, D. A. in *Mars* (eds Kieffer, H. H., Jakosky, B. M., Snyder, C. W. & Matthews, M. S.) 934–968 (Univ. Arizona Press, Tucson, 1992).
18. Martinez, I., Agrinier, P., Schärer, U. & Jovoy, M. *Earth planet. Sci. Lett.* **121**, 559–574 (1994).
19. Ash, R. A., Knott, S. F. & Turner, G. *Meteoritics* **30**, 485 (1995).
20. Baker, V. R., Carr, M. H., Gulick, V. C., Williams, C. R. & Marley, M. S. in *Mars* (eds Kieffer, H. H., Jakosky, B. M., Snyder, C. W. & Matthews, M. S.) 493–522 (Univ. Arizona Press, Tucson, 1992).
21. Ridley, W. I., Plumlee, G. S., DeBraal, J. D. & Reed, M. H. *Meteoritics* **30**, 565 (1995).
22. Griffith, L. L. & Shock, E. L. *Nature* **377**, 406–403 (1995).

ACKNOWLEDGEMENTS. We thank L. Anovitz for helpful discussions, and A. Treiman, W. M. Murphy and O. Vidal for reviews. This work was supported by NASA (H.Y.M.) and NSF (R.P.H.).

CORRESPONDENCE should be addressed to R.P.H. (e-mail rph@po.cwru.edu).

Spectroscopic evidence for a pseudogap in the normal state of underdoped high- T_c superconductors

H. Ding^{*†}, T. Yokoya[‡], J. C. Campuzano^{*†},
T. Takahashi[‡], M. Randeria[§], M. R. Norman[†],
T. Mochiku[¶], K. Kadowaki[¶] & J. Giapintzakis[#]

^{*} Department of Physics, University of Illinois at Chicago, Chicago, Illinois 60607, USA

[†] Materials Science Division, Argonne National Laboratory, Argonne, Illinois 60439, USA

[‡] Department of Physics, Tohoku University, Sendai 980, Japan

[§] Tata Institute of Fundamental Research, Bombay 400005, India

[¶] Institute of Materials Science, University of Tsukuba, Tsukuba, Ibaraki 305, Japan

[¶] National Research Institute for Metals, Tsukuba, Ibaraki 305, Japan

[#] Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

It is well known that BCS mean-field theory is remarkably successful in describing conventional superconductors. A central concept of BCS theory is the energy gap in the electronic excitation spectrum below the superconducting transition temperature, T_c . The gap also serves as the order parameter: quite generally, long-range phase coherence and a non-zero gap go hand-in-

hand¹. But in underdoped high- T_c superconductors there is considerable evidence that a pseudogap (a suppression of spectral weight) is already formed in the normal state above T_c —first, from studies of the spin excitation spectrum^{2–5,24}, which measure a ‘spin gap’, and later from a variety of other probes^{6–10}. Here we present a study of underdoped Bi₂Sr₂CaCu₂O_{8+δ} (Bi2212) using angle-resolved photoemission spectroscopy (ARPES), which directly measures the momentum-resolved electron excitation spectrum of the CuO₂ planes. We find that a pseudogap with d -wave symmetry opens up in the normal state below a temperature $T^* > T_c$, and develops into the d -wave superconducting gap once phase coherence is established below T_c .

In ARPES, photons incident upon a sample cause electrons to be ejected, whose energies and momenta are measured. From this one obtains the electronic excitation spectrum of the sample. Our experiments were carried out at the Synchrotron Radiation Center, Wisconsin, following procedures described previously¹¹, with 19- or 22-eV photons at an energy resolution of 22–27 meV (full-width at half-maximum, FWHM) and an angular window of $\pm 1^\circ$. Sample underdoping was achieved by adjusting the oxygen partial pressure during annealing of the float-zone grown crystals. These crystals have sharp X-ray diffraction rocking curves with structural coherence lengths of 1,250 Å, similar to the near optimally doped samples studied earlier. This high degree of order apparently also helps in maintaining the structural stability of the sample in an ultra-high vacuum of $< 4 \times 10^{-11}$ torr under constant-temperature cycling, even up to 300 K, without measurable sample degradation. All samples show a very flat surface after cleaving, which is essential for determining the intrinsic momentum ($\mathbf{k}$) dependence of the gap.

We have studied several samples ranging from overdoped to underdoped, where optimal doping corresponds to a T_c of 92 K. In this Letter we will focus on a moderately underdoped sample with a T_c of 83 K (transition width 2 K), and a heavily underdoped sample with a T_c of 10 K (width >5 K), and contrast these results with a near-optimal (slightly overdoped) 87 K sample (width 1 K). We will simply label the samples by their onset T_c s.

It is important to separate the effects of changes in carrier concentration from those due to disorder. Because the superconducting gap for near-optimal Bi2212 is consistent with a d -wave order parameter^{12–14}, for which disorder leads to T_c reduction, we must ensure that the samples we call underdoped have reduced T_c due to lower doping, and not due to increased disorder. In Fig. 1 we compare the spectra for the near-optimal (87 K) and underdoped samples with an irradiated sample, whose T_c has been brought down to 79 K by intentionally disordering an 87 K material by high-energy electron irradiation. We see that the 83 K underdoped sample has a much broader normal-state spectrum than the irradiated sample with a similar T_c (Fig. 1a), and the 10 K sample has an even broader spectrum. Given the considerable evidence linking ARPES linewidths with electron–electron interactions^{15,16}, this suggests that underdoped systems are more strongly correlated.

In fact, the spectra of the underdoped sample are so broad that it is difficult to define a spectral peak above T_c . It is therefore significant to note in Fig. 1b that for $T \ll T_c$ the line shape sharpens up and a well defined peak is seen for the 83 K sample (at those $\mathbf{k}$ -points at which there is a sizeable superconducting gap). From Fig. 1b it is clear that all samples in the superconducting state have very similar low-temperature lineshapes, even though their lineshapes differ significantly in the normal state. We note that the 10 K sample is still in the normal state at $T = 14$ K, and therefore the linewidth sharpening¹⁵ is associated with superconductivity, rather than with the temperature. Because underdoped samples have broad line shapes which are difficult to fit by any simple spectral function¹¹, we use the shift of the midpoint of the leading edge of a spectrum from the zero of the binding energy (that is, the Fermi energy) as an estimate of the gap. We expect, based on our earlier studies¹⁴, that this estimate is qualitatively correct for both the magnitude and the $\mathbf{k}$ -dependence of the gap.

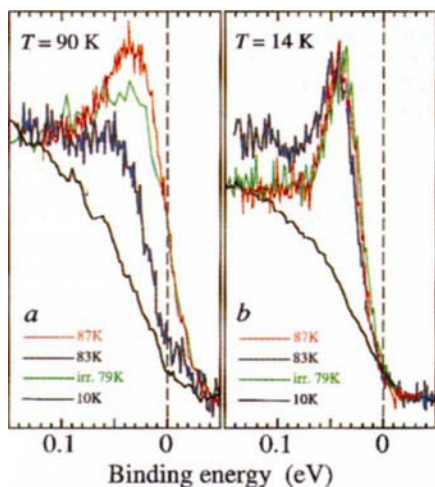


FIG. 1 ARPES spectra at the $(\pi,0)$ – (π,π) Fermi surface crossing point for four samples (labelled by T_c): near-optimal (87 K, red trace), underdoped (83 K and 10 K, blue trace and black trace, respectively), and irradiated (79 K, green trace), at $T = 90$ K (a), and $T = 14$ K (b). Note that all superconducting-state spectra are similar, while the normal-state spectra differ significantly. A comparison of linewidths between the underdoped 83 K and irradiated 79 K samples suggests that disorder is not the main cause of linewidth broadening.

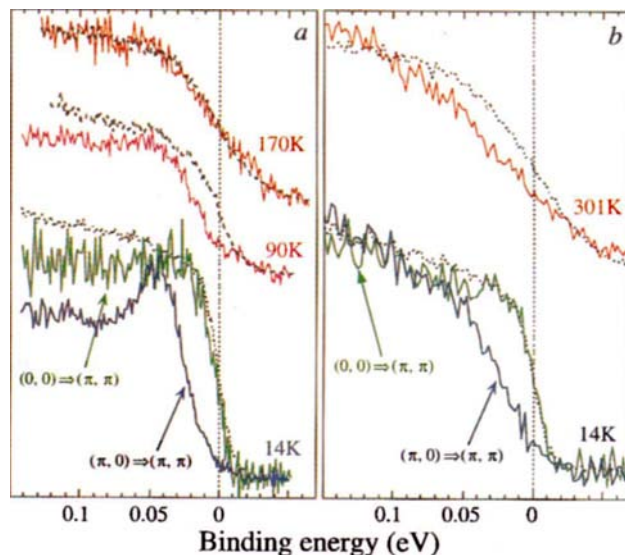


FIG. 2 Spectra (solid lines) for underdoped samples at the $(\pi,0)$ – (π,π) Fermi surface crossing, along with reference Pt spectra (broken lines) at different temperatures. a, Spectra (22-eV photons) of the 83 K sample. At 14 K there is a gap in the superconducting state, which persists into the normal state in the 90 K spectrum, eventually closing at $T^* = 170$ K. b, Spectra (19-eV photons) of the 10 K sample. Note a pseudogap at the $(\pi,0)$ – (π,π) Fermi surface crossing exists in the normal state from 14 to 301 K. For both a and b we also plot the spectrum at the $(\pi,0)$ – (π,π) Fermi surface crossing at 14 K, showing a vanishing gap.

Spectra for the 83 K sample at three different temperatures are shown in Fig. 2a together with spectra (broken lines) from a polycrystalline Pt foil in electrical contact with the sample. The Pt leading-edge midpoint determines the Fermi energy. At the $\bar{M} = (\pi,0)$ to $\bar{Y} = (\pi,\pi)$ Fermi surface crossing, we find that the leading edge of Bi2212 is pushed to positive binding energies relative to Pt at 14 K indicating a large superconducting gap. This gap does not fully close above T_c , as seen in the 90 K data, and we will call this a ‘pseudogap’. We use this term because the linewidths above T_c are so broad that there is non-zero spectral weight at the Fermi energy, so that a true gap does not exist. Similar results have also been recently reported in refs 17 and 18. The pseudogap persists up to 170 K, which we identify as T^* , where no visible shift is seen between the Bi2212 and Pt leading edges and the pseudogap vanishes. We contrast this behaviour to that along the (π,π) direction, where no gap is seen even well below T_c .

Above $T^* = 170$ K, the Fermi surface of the 83 K sample is recovered, by which we mean that we can experimentally map out a closed contour of gapless excitations in $\mathbf{k}$ -space. Below T^* if one makes a sequence of cuts in $\mathbf{k}$ -space normal to the Fermi surface, then those points where the gap is smallest along a cut define the minimum gap locus. For the 83 K sample, this locus is found to coincide with the Fermi surface measured above T^* , and is a large hole-like barrel enclosing the (π,π) point. Preliminary results indicate a slight reduction in the enclosed area relative to optimal doping; details will be presented in a future publication.

In Fig. 2b we show similar results for the 10 K sample. Again, a large pseudogap in the normal state is seen along $\bar{M}\bar{Y}$ and no gap along $\bar{\Gamma}\bar{Y}$ ($r = (0,0)$). For this sample, though, the large pseudogap does not close even at 301 K, the highest temperature at which we did measurements. Thus $T^* > 301$ K at this doping level. As a gapless Fermi surface is not recovered even at 301 K, we use the minimum gap locus (defined above) to infer the shape and size of the Fermi surface which would presumably be recovered above T^* . The minimum gap locus for the 10 K sample is a large barrel, as in the optimal and lightly-underdoped cases. Preliminary results indicate that its volume is smaller for the 10 K sample

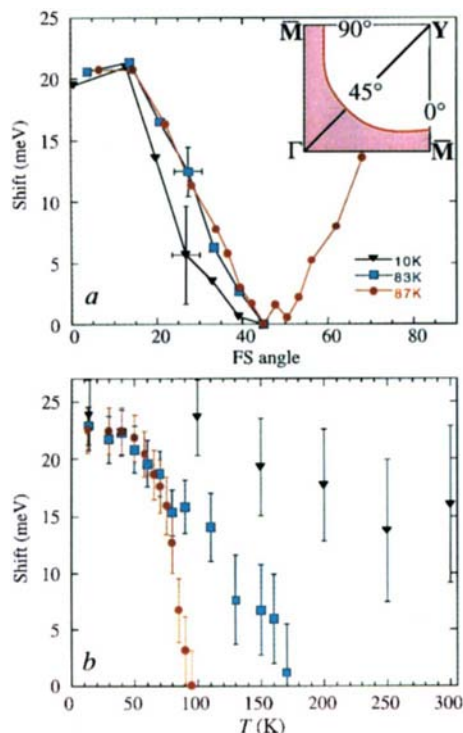


FIG. 3 Momentum- and temperature-dependence of the gap estimated from leading-edge shift (see text). *a*, $\mathbf{k}$ -dependence of the gap in the 87 K T_c , 83 K T_c and 10 K T_c samples, measured at 14 K. The inset shows the Brillouin zone with a large Fermi surface (FS) closing the (π, π) point, with the occupied region shaded. Note the striking similarity in both the magnitude and the $\mathbf{k}$ -dependence of the gap in the three samples with widely differing T_c s. *b*, Temperature dependence of the maximum gap in a near-optimal 87 K sample (circles), underdoped 83 K (squares) and 10 K (triangles) samples. Note smooth evolution of the 'gap' from superconducting to normal state for the 83 K sample.

than the 83 K sample, as expected. Note that the existence of a large Fermi surface closed about (π, π) would contradict several theories of lightly doped Mott insulators which predict small hole pockets centred above $(\pi/2, \pi/2)$.

In Fig. 3a we plot the angular variation of the leading-edge shifts measured from the Fermi energy at $T = 14$ K along the Fermi surface (FS). For the sake of comparison between samples, we vertically offset the shifts in Fig. 3a so that the shift at 45° is zero. (The offset is -3 meV for the 83 K sample and $+2$ meV for the 10 K sample.) We have seen similar sample-to-sample variations of the offsets for optimal-doped samples with the same T_c , so we do not regard these offsets as significant. Note the large error bars on the results for 10 K sample, as this has a rather broad spectrum at 14 K which makes it difficult to accurately determine the midpoint of its leading edge. To address the pseudogap anisotropy for the 10 K sample, we have measured its angular dependence on the locus of minimum gap just as we did for the 83 K sample.

Remarkably, both the magnitude and the $\mathbf{k}$ -dependence of the gap are similar for the three samples, and independent of carrier concentration, as seen from Fig. 3a, even though T_c and T^* change appreciably. The doping-independence of the pseudogap magnitude has also been suggested in ref. 19. We note that the observed shift in the 87 K sample corresponds to a superconducting gap with a maximum value of 33 meV as determined by a spectral function analysis of the data^{11,14}. The $\mathbf{k}$ -dependence of the gap in all the samples is very similar to that of the near-optimal

87 K material, which has been studied in detail by our group^{13,14} and found to be consistent with a d -wave $|\cos k_x - \cos k_y|$ gap, as suggested in earlier ARPES studies¹².

We emphasize that, as far as ARPES is concerned, there is no difference between the gap seen above and below T_c , except in one respect which we discuss below. This is clearly seen for the 83 K sample in Fig. 3b, where we plot the temperature dependence of the leading-edge shifts, without any offsets, at the $M = (\pi, 0)$ to $Y = (\pi, \pi)$ Fermi surface crossing for various samples. (Fig. 3a and Fig. 3b were obtained from different data sets, and therefore show slightly different values of the maximum gap). For the 83 K sample there is no signature of the phase transition at T_c in the magnitude of the gap, which eventually vanishes at $T^* = 170$ K. In contrast, in the near-optimal 87 K sample T^* (at which the gap closes) is close to T_c . As for the 10 K sample, a pseudogap regime exists over the entire range of temperatures studied. The one difference between the spectra in the superconducting and the pseudogap states is in their linewidths. Above T_c the linewidth is very broad, and spectral weight is present at the Fermi energy even though the leading edge is shifted due to the pseudogap. In contrast, well below T_c , the linewidth is resolution-limited, and a clear gap is observed, with very small spectral weight at the Fermi energy due to thermally excited quasi-particles.

These observations are remarkable and, to our knowledge, such effects have not been seen before in ordinary bulk superconductors. Part of our initial motivation was the prediction^{20,21} in related model systems that the pseudogap is a normal-state precursor of the superconducting gap above T_c , where there is a pairing amplitude without phase coherence²². At present, however, there is no quantitative theory to explain our observations of the evolution of the planar Fermi surface, T^* , T_c and the gap as functions of the carrier concentration.

We summarize our results in the form of a phase diagram for Bi2212 in Fig. 4. At present, this is schematic, as we do not know the precise carrier concentrations for our samples, except for optimal doping, which corresponds approximately to 0.17 holes per Cu (ref. 16). (For hole concentrations in polycrystalline samples, see ref. 23.) The filled symbols denote the superconducting T_c determined by our susceptibility measurements. The full

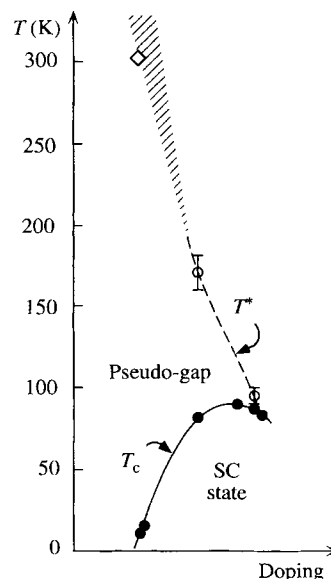


FIG. 4 Schematic phase diagram of Bi2212 as a function of doping. The filled symbols are the measured T_c s for the superconducting phase transition from magnetic susceptibility. The open symbols are the T^* at which the pseudogap closes; for the $T_c = 10$ K sample, the symbol at 301 K is a lower bound on T^* (thus the hatching). The region between T^* and T_c is an unusual 'normal' state with a pseudogap in the electronic excitation spectrum.

line (which is a guide to the eye) through these points represents a phase boundary between the superconducting (SC) and 'normal' phases. The latter exhibits a pseudogap in its electronic excitation spectrum, which closes only at T^* . The experimental points for T^* are denoted by open symbols, and the dashed line through these points denotes a crossover line above which a Fermi surface of gapless excitations is recovered. We find that T^* increases with underdoping, consistent with earlier results^{9,10}, and 301 K represents only a lower bound on the T^* for our $T_c = 10$ K sample, in which T_c is lower by a factor of at least 30 relative to T^* . Given their same $\mathbf{k}$ -dependence above and below T_c , and smooth evolution through T_c , our results suggest that the normal-state pseudo-gap is closely related to the superconducting gap below T_c . □

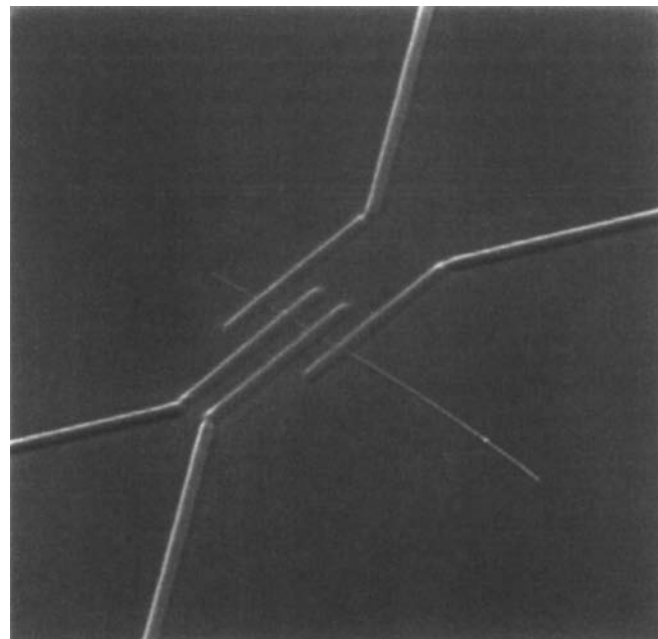


FIG. 1 A single nanotube connected to four 80-nm wide tungsten wires and imaged by FIB.

Received 15 April; accepted 11 June 1996.

1. Schrieffer, J. R. *Theory of Superconductivity* (Cummings, Reading, 1964).
2. Warren, W. W. Jr et al. *Phys. Rev. Lett.* **62**, 1193–1196 (1989).
3. Takigawa, M. et al. *Phys. Rev.* **B43**, 247–257 (1991).
4. Alloul, H., Mahajan, A., Casalta, H. & Klein, O. *Phys. Rev. Lett.* **70**, 1171–1174 (1993).
5. Rossat-Mignod, J. et al. *Physica B* **186–188**, 1–8 (1993).
6. Homes, C. C., Timusk, T., Liang, R., Bonn, D. A. & Hardy, W. N. *Phys. Rev. Lett.* **71**, 1645–1648 (1993).
7. Basov, D. N., Mook, H. A., Dabrowski, B. & Timusk, T. *Phys. Rev.* **B52**, 13141–13144 (1995).
8. Loram, J. W., Mirza, K. A., Cooper, J. R. & Liang, W. Y. *Phys. Rev. Lett.* **71**, 1740–1743 (1993).
9. Tallon, J. L., Cooper, J. R., de Silva, P., Williams, G. V. M. & Loram, J. W. *Phys. Rev. Lett.* **75**, 4114–4117 (1995).
10. Batlogg, B. et al. *Physica C* **235–240**, 130–133 (1994).
11. Ding, H. et al. *Phys. Rev. Lett.* **74**, 2784–2787 (1995); **75**, 1425 (1995).
12. Shen, Z.-X. et al. *Phys. Rev. Lett.* **70**, 1553–1556 (1993).
13. Norman, M. R., Randeria, M., Ding, H., Campuzano, J. C. & Bellman, A. F. *Phys. Rev.* **B52**, 15107–15110 (1995).
14. Ding, H. et al. <http://xxx.lanl.gov/archive/cond-mat/9603044>.
15. Randeria, M. et al. *Phys. Rev. Lett.* **74**, 4951–4954 (1995).
16. Ding, H. et al. *Phys. Rev. Lett.* **76**, 1533–1536 (1996).
17. Loeser, A. G., Shen, Z.-X. & Dessau, D. S. *Physica C* **263**, 208–213 (1996).
18. Loeser, A. G. *Science* (in the press).
19. Loram, J. W., Mirza, K. A., Wade, J. M., Cooper, J. R. & Liang, W. Y. *Physica C* **235–240**, 134–137 (1994).
20. Randeria, M., Trivedi, N., Moreo, A. & Scalettar, R. T. *Phys. Rev. Lett.* **69**, 2001–2004 (1992).
21. Trivedi, N. & Randeria, M. *Phys. Rev. Lett.* **75**, 312–315 (1995).
22. Emery, V. & Kivelson, S. A. *Nature* **374**, 434–437 (1995).
23. Groen, W. A., de Leeuw, D. M. & Feiner, L. F. *Physica C* **165**, 55–61 (1990).
24. Imai, J. et al. *Physica C* **162–164**, 169–170 (1989).

ACKNOWLEDGEMENTS. This work was supported by the US NSF, the US Dept of Energy, Basic Energy Sciences, the US NSF Science and Technology Center for Superconductivity, the Japan Society for the Promotion of Science, NEDO, and the Ministry of Education, Science and Culture of Japan. The Synchrotron Radiation Center is supported by the NSF.

CORRESPONDENCE should be addressed to J.C.C. (e-mail: jcc@uicwsc.phy.uic.edu).

Electrical conductivity of individual carbon nanotubes

T. W. Ebbesen*, H. J. Lezec†, H. Hiura‡, J. W. Bennett*, H. F. Ghaemi* & T. Thio*

* NEC Research Institute, 4 Independence Way, Princeton, New Jersey 08540, USA

† Micrion Europe GmbH, Garmischer Strasse 4/V, 80339 München, Germany

‡ Fundamental Research Laboratories, NEC, 34 Miyukigaoka, Tsukuba 305, Japan

THE interest in carbon nanotubes has been greatly stimulated by theoretical predictions that their electronic properties are strongly modulated by small structural variations^{1–8}. In particular, the diameter and the helicity of carbon atoms in the nanotube shell are believed to determine whether the nanotube is metallic or a semiconductor. Because of the enormous technical challenge of making measurements on individual nanotubes, however, experimental studies have been limited mainly to bulk measurements⁹, which indicate only that a fraction of the nanotubes are metallic or narrow-band semiconductors¹⁰. Recently, measurements of the magneto-conductance of a single multi-shell

nanotube in a two-probe configuration showed that the transport is characterized by disorder and localization phenomena¹¹. To avoid possible ambiguities due to poor sample contacts, four-probe measurements are needed. Here we report four-probe measurements on single nanotubes made by lithographic deposition of tungsten leads across the tubes. We find that each multi-shell nanotube has unique conductivity properties. Both metallic and non-metallic behaviour are observed, as well as abrupt jumps in conductivity as the temperature is varied. The differences between the electrical properties of different nanotubes are far greater than expected. Our results suggest that differences in geometry play a profound part in determining the electronic behaviour.

The carbon nanotubes were prepared by the carbon-arc method¹² which yields nanotubes with the highest degree of structural integrity and then further annealed at 3,120 K for 15 min under argon. This high-temperature treatment eliminates many defects as seen by electron spin resonance spectroscopy¹⁰. The nanotubes were spread on an oxidized silicon surface between gold pads. Using a Micrion focused-ion-beam (FIB) System 9100 (30 keV Ga ions, 10 nm nominal spot diameter), the surface was visualized using a low beam current (4 pA) in search of a nanotube with the right length (~2–4 µm). The total dose during this imaging was 2×10^{14} ions per cm² (that is, two ions per 100 Å²). Once found, four 80-nm-wide tungsten leads were patterned on the nanotube by ion-induced deposition of tungsten from a W(CO)₆ carrier gas. The patterning is performed under computer control without further visualizing the sample. All the samples were prepared with exactly the same procedure and doses. An example of the resulting connections is shown in Fig. 1. These small leads were then connected to four surrounding gold pads. The distance between the contacts on the nanotube is in all cases between 0.3 and 1.0 µm. It was verified that with such spacings the resistance between the leads (without the nanotube) was larger than the limit of our measurement, 10 GΩ. The diameter of each nanotube was measured with a Nanoscope III atomic force microscope. The four-probe resistances were measured in d.c. using a Keithley 220 current supply and either a Keithley 196 multimeter or a Keithley 617 electrometer, depending on the sample resistances. The applied voltage was held below 1 V at all

times. Comparison with two-probe measurements show that the nanotube-tungsten contact resistances are small ($\sim 5 \text{ k}\Omega$) in the highly conductive tubes but there is evidence for the formation of a Schottky barrier in the more insulating nanotubes. In order to measure the temperature and magnetic-field effects, the samples were mounted on a temperature-controlled sample holder and placed in the field of a superconducting magnet.

The nanotube can easily be damaged by even the smallest static discharges, which is not surprising considering their very small dimensions. Many samples were damaged in this way before further measurements could be made despite numerous precautions. Therefore complete four-probe temperature and magnetic-field dependences were not obtained on all the nanotubes; nevertheless, enough results were collected to give a sense of the diversity in nanotube electronic properties. It should be noted that the current-carrying capacity of the nanotubes is extremely high. For metallic tubes, current densities as high as $6 \times 10^6 \text{ A cm}^{-2}$ were reached without damaging the sample.

Table 1 shows the room temperature four-probe resistances of eight nanotubes of various diameters. These results already indicate that the nanotube resistances vary by orders of magnitude. The resistances we are measuring are probably those of several shells; because it is not known how the current flow in the various layers is weighted, it is difficult to convert the resistances into meaningful resistivities. For purpose of comparison only, the last two columns of the tubes give the resistance per unit length and the resistivity in terms of a three-dimensional conductor. The in-plane resistivity of crystalline graphite is $\sim 3.8 \times 10^{-5} \Omega \text{ cm}$ (ref. 13); Table 1 suggests that nanotubes can reach this lower limit.

The temperature dependence of the conductivity of various nanotubes is shown in Fig. 2. First, we note that Fig. 2a–c (which shows four-probe measurements) reveals great differences between nanotubes (NT). The resistivity of NT8 decreases linearly by only $\sim 10\%$ between 4 and 300 K. NT7a shows a large transition with a drop in resistance as T is lowered past 220 K, while NT6 has a reverse transition with the resistance becoming unmeasurably high for $T < 200 \text{ K}$. Unlike all the other nanotubes analysed, NT6 showed significant contact resistance and negative voltage at certain temperatures. This is typical of a filamentous current and indicates a complex conductive path inside the nanotube.

Measurements on other nanotubes, or segments of nanotubes, on which we were able to do two-probe measurements, are shown in Fig. 2d. The temperature dependences of NT2 and NT1 are similar to NT8, except that NT1 has a transition just below 250 K. When comparing four-probe and two-probe measurements on the same nanotube, it became clear that sometimes different segments of a nanotube do not have the same temperature profile. Figure 2b and d illustrate such differences between the centre section and a side section of NT7, respectively.

The nanotubes studied fall into broad categories. First of all, if we exclude the nanotubes having large transitions (NT6 and NT7a), nanotubes 1, 2, 7 and 8 show a very consistent percentage increase in resistivity with decreasing temperature. The change is $\sim 10\%$ from 300 to 4 K and is close to linear. A fit to a thermally activated temperature-dependence yields a gap of $< 0.1 \text{ meV}$ which is physically meaningless. For all practical purposes, these nanotubes are gapless and the temperature dependence is probably of different origin. However, at the same time, these nanotubes show significant variation in resistivities (Table 1).

TABLE 1 Four-probe measurement of the conductivity of various nanotubes at room temperature

Nanotube no.	Radius* (nm)	Pitch† (μm)	Resistance (Ω)	Resistance per μm ($\Omega \mu\text{m}^{-1}$)	Resistivities‡ ($\Omega \text{ cm}$)
0	5.0	1.0	$\geq 10^8$	$\geq 10^8$	≥ 0.8
1§	10.2	0.3	10.8×10^3	3.6×10^4	1.2×10^{-4}
2	~ 3.0	0.35	10.5×10^3	3.0×10^4	7.5×10^{-5}
4	6.3	0.5	2.4×10^8	4.8×10^8	5.8
5	3.6	0.9	$\geq 10^6$	$\geq 10^6$	$> 4 \times 10^{-3}$
6	9.1	1.0	2.0×10^2	2.0×10^2	5.1×10^{-6}
7	6.1	0.5	4.3×10^4	8.6×10^4	9.8×10^{-4}
8	7.4	0.5	6.0×10^3	1.2×10^4	2.0×10^{-4}

* Within $\pm 10\%$.

† Distance between leads probing the sample.

‡ Resistivities estimated assuming a three-dimensional conductor with an average inner hollow of 2 nm (ref. 9).

§ Non-annealed nanotube.

Considering both the gapless nature and the low values of the resistivities, these nanotubes can be considered to be metallic. By analogy with carbon fibres¹⁴, the variation in resistivities and their temperature dependence are possibly due to the interplay of changes in carrier concentration and mobilities in the metallic tubes. Finally, when magnetic fields of up to 12 T were applied perpendicular to the long axis of the tubes, the effect was negligible. Only a slight positive magneto-resistance was observed (for example, $\sim 1\%$ for NT7b (Fig. 2f) at $T = 7 \text{ K}$ and $H = 10 \text{ T}$).

Nanotubes 0, 4 and 5 in Table 1 cannot be considered metallic: their resistivities are rather high, but more significantly the tubes show a strong decrease in resistance with increasing temperature, indicating a thermally activated process. Fig. 3 shows a standard $\log R$ (where R is the resistance) versus $1/T$ plot for these nanotubes. While the profiles are typical of semiconductors, and the slopes yield activation energies of 0.1 and 0.3 eV for NT4 and NT5 respectively, the lack of linearity in NT4

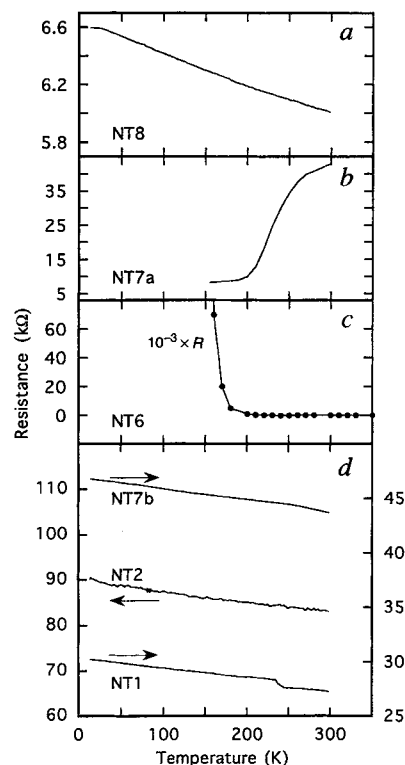


FIG. 2 Temperature dependence of the resistance of different nanotubes: a–c are four-probe measurements; d shows two-probe results.

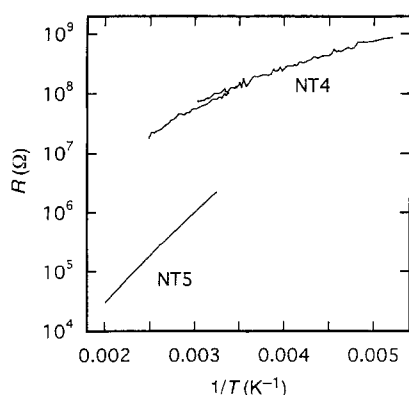


FIG. 3 Semilogarithmic plot of the resistance of nanotubes 4 and 5 (see Table 1) versus the inverse of the temperature (traces labelled NT4 and NT5 are four-probe and two-probe results, respectively).

suggests a more complex behaviour. Further analyses of more tubes are needed to clarify this temperature dependence.

Some nanotubes show large transitions with temperature and they may occur in either direction. The sudden increase in resistance with decreasing temperature in NT6 and NT1 might be due to a Peierls transition, while the profile of NT7a could be due to an insulator-to-metal transition. Further investigation of such transitions will be necessary to establish their exact identity. These transitions may be the property of either one or more shells in the tube and/or a segment of the tube as the comparison of Fig. 2b and f clearly illustrates. Therefore the observed conductivity of a single nanotube, even by a four-probe measurement, may sometimes be the sum of various segments in series between the probes.

Langer *et al.*¹¹ have studied the conductance of a non-annealed 20-nm nanotube in a two-probe configuration. In contrast to the slight positive magneto-resistance and small linear temperature dependence in our samples, they reported a strong negative magneto-resistance and a logarithmic fall in resistance ($R(T) \approx -\log(T)$) with increasing temperature. Their results were attributed to weak localization in a disordered two-dimensional conductor¹¹. The difference could be simply be part of the wide spectrum of nanotube types. It is already well established that the bulk properties of nanotubes vary significantly depending on where they were made, which of course reflects a variation of the distribution of the types of individual tubes composing the bulk^{9,10}. The difference could also stem from the methodology used in preparing the tube for transport measurements. Since submission of this paper, another two-probe study of nanotubes has been published¹⁵. Comparison of the measured resistances with our results is difficult because no temperature or magnetic field dependence were reported and the nanotubes were produced catalytically.

The presence of both metallic and non-metallic nanotubes in these samples is consistent with theoretical predictions that single nanotubes may be metallic or semiconducting depending on their helicity²⁻⁵, and that this should be observable for nanotubes with diameters up to 30 nm for the wide-gap semiconducting types⁵. In multi-layer nanotubes such as those studied here, the metallic and semiconducting nature of the tubes are expected to be modified by the interlayer interactions^{7,16}. Furthermore, we cannot unequivocally rule out the possibility that structural defects do not contribute to the observed behavior, even though we used annealed arc-produced nanotubes. Such factors probably contribute to the large variations observed in the resistivity between the metallic tubes. Combined with the appearance of abrupt changes in resistivity with temperature, it clearly shows that nanotubes have a far wider spectrum of electronic properties than had been

expected. Each nanotube is unique, demonstrating the strong correlation of structure and electronic properties at this scale. □

Received 2 April; accepted 11 June 1996.

1. Mintmire, J. W., Dunlap, B. I. & White, C. T. *Phys. Rev. Lett.* **68**, 631–634 (1992).
2. Hamada, N., Sawada, A. & Oshiyama, A. *Phys. Rev. Lett.* **68**, 1579–1581 (1992).
3. Saito, R., Fujita, M., Dresselhaus, G. & Dresselhaus, M. S. *Appl. Phys. Lett.* **60**, 2204–2206 (1992).
4. Tanaka, K., Okahara, K., Okada, M. & Yamabe, T. *Chem. Phys. Lett.* **191**, 469–472 (1992).
5. White, C. T., Robertson, D. H. & Mintmire, J. W. *Phys. Rev. B* **47**, 5485–5488 (1993).
6. Jishi, R. A., Dresselhaus, M. S. & Dresselhaus, G. *Phys. Rev. B* **48**, 11385–11389 (1993).
7. Lambin, Ph., Philippe, L., Charlier, J. C. & Michenaud, J. P. *Computat. Mater. Sci.* **2**, 350–356 (1994).
8. Blase, X., Benedict, L. X., Shirley, E. L. & Louie, S. G. *Phys. Rev. Lett.* **72**, 1878–1881 (1994).
9. Ebbesen, T. W. A. *Rev. Mater. Sci.* **24**, 235–264 (1994).
10. Kosaka, M., Ebbesen, T. W., Hiura, H. & Tanigaki, K. *Chem. Phys. Lett.* **233**, 47–51 (1995).
11. Langer, L. *et al. Phys. Rev. Lett.* **76**, 479–482 (1996).
12. Ebbesen, T. W. & Ajayan, P. M. *Nature* **358**, 220–222 (1992).
13. Primak, W. & Fuchs, L. H. *Phys. Rev.* **95**, 22–30 (1954).
14. Dresselhaus, M. S., Dresselhaus, G., Sugihara, K., Spain, I. L. & Goldberg, H. A. in *Graphite Fibers and Filaments* (eds Gonser, U., Mooradian, A., Muller, K. A., Panish, M. B. & Sakaki, H.) 173–229 (Springer, New York, 1988).
15. Dai, H., Wong, E. W. & Lieber, C. M. *Science* **272**, 523–526 (1996).
16. Benedict, L. X., Crespi, V. H., Louie, S. G. & Cohen, M. L. *Phys. Rev. B* **52**, 14935–14939 (1995).

ACKNOWLEDGEMENTS. We thank D. J. Chadi and N. S. Wingreen for their comments on these results, and A. Dziesiaty, W. Muth, H. Gieser and H. Zimmermann for their assistance.

CORRESPONDENCE should be addressed to T.W.E. and H.J.L.

El Niño-like climate change in a model with increased atmospheric CO₂ concentrations

Gerald A. Meehl & Warren M. Washington

National Center for Atmospheric Research, Boulder, Colorado 80307-3000, USA

SEA surface temperatures in the tropical Pacific Ocean increased on average by several tenths of a degree during the 1980s and early 1990s¹⁻⁴, contributing to the observed global warming during this period⁵. Here we investigate the possible causes of this Pacific warming, using a global coupled ocean–atmosphere general circulation model incorporating increasing concentrations of atmospheric carbon dioxide. In the model, cloud cover and cloud albedo feedbacks contribute to tropical Pacific sea surface temperature increases that are greater east of 180° longitude, with attendant shifts in large-scale precipitation patterns and mid-latitude circulation anomalies in the north Pacific. These anomalies resemble some aspects of El Niño events, as well as features associated with recent observed Pacific-region climate anomalies. The resemblance to El Niño complicates the problem of detection and attribution of climate change, and suggests that depletion of freshwater resources⁶ may be an additional hazard of greenhouse warming for populations in the western Pacific region.

A second-generation global coupled general circulation climate model was integrated for 75 years with atmospheric CO₂ levels increasing at a rate of 1% per year compounded. The last 20 years of this experiment are analysed (years 56–75; the CO₂ level doubles at approximately year 70) and compared to a control integration averaged over that same period with present-day amounts of CO₂. No flux corrections are used at the air–sea interface in the coupled model. The observed area-averaged sea surface temperature (SST) difference, western minus eastern Pacific area (as defined in Table 1), is 4.19 °C, and the coupled model difference is 3.21 °C. Thus the model is able to maintain a

reasonable longitudinal SST gradient across the Pacific albeit with mean SSTs somewhat warmer than observed (Table 1; also Fig. 3a in ref. 7).

A simple cloud albedo feedback parametrization and a mass flux convective scheme are included in the atmospheric model. The latter represents the super greenhouse effect in the model^{7,8}. The former parametrization accounts for the observed relationship between very warm SSTs, deep convection and bright

clouds^{8,9} such that if SST is >303 K and deep convection is taking place at that gridpoint, middle- and upper-level convective cloud albedos (starting at 0.3 and 0.15, respectively) increase linearly to 0.6 for the SST range of 303–308 K. Though the total greenhouse effect increases rapidly for SST values above ~ 300 K (ref. 8) thus suggesting an active cloud albedo feedback for SST above 300 K, we use the more conservative SST threshold of 303 K for our cloud albedo feedback parametrization. This SST

TABLE 1 Area averages from the coupled model

	SST (°C)	Abs. Solar (W m ⁻²)	Net IR (W m ⁻²)	Sens. (W m ⁻²)	Latent (W m ⁻²)	Adv. (W m ⁻²)	Q-con. (%)	Q (%)	TCLD (%)	Alb. (%)
E. Pacific	+3.49	+2.23	+0.92	+2.11	+8.67	+7.65	+62	+22	-13	-4
W. Pacific	+2.21	-3.92	+0.28	+0.88	+0.81	+5.28	+3	+16	-5	+1

The western tropical Pacific area is defined for 5° S–5° N, 140° E–170° E, and the eastern tropical Pacific area is defined for 5° S–5° N, 120° W–90° W. Computed minus observed³³ area averaged SST differences for the control case are 1.89 °C for the eastern Pacific area, and 0.91 °C for the western Pacific area, with the coupled model somewhat systematically warmer than observed as noted in the text and elsewhere⁷. To compare the effects of the cloud albedo feedback on the longitudinal SST response across the tropical Pacific to an increase of CO₂, we note that the eastern Pacific area-averaged SSTs in the first column warm 58% more than the western Pacific area-averaged SSTs. In contrast, the eastern Pacific area warms only 30% more than the western Pacific area in a different version of the coupled model without cloud albedo feedback³⁴. In a version of the atmospheric model coupled to a non-dynamic slab ocean⁹ the east Pacific area warmed 20% more than the western Pacific area when cloud albedo feedback was included, but only 5% more without cloud albedo feedback. Using a sample from April (following Fig. 1) from the present coupled model both with and without the cloud albedo scheme, the change in total absorbed solar radiation at the top of the atmosphere for the western Pacific area with cloud albedo feedback is -5.1 W m⁻². The contribution due to changes in cloud albedo alone from the cloud albedo feedback scheme (the difference in C_s at the top of the atmosphere with and without the cloud albedo feedback scheme) is -11.9 W m⁻². Thus the contribution due to cloud decreases is $+6.8$ W m⁻². In the eastern Pacific area, decreases in cloud fraction (increased convective cloud with smaller grid fraction coverage) contribute to increased absorbed solar radiation ($+13.8$ W m⁻²), but there is also a contribution from increased cloud albedo of -7.2 W m⁻² (the difference in C_s with and without the cloud albedo feedback scheme). Thus the contribution due to cloud changes is $+21.0$ W m⁻². There is 65% more solar radiation reflected by the clouds in the west compared to the east solely due to the cloud albedo feedback in the model, but there are also contributions from changes in cloud amount and type in response to the dynamical coupling between ocean and atmosphere. Without cloud albedo feedback and ocean dynamics, the major contributor to decreasing east–west SST gradient is temperature–evaporation feedback that can produce similar El Niño-like climate-change signals associated with tropical Pacific warming^{34,35}. Shown in this table are area-averaged results for increased CO₂ experiment minus control for the present version of the coupled model, December–January–February, 20-year averages. Total cloud and albedo for the eastern Pacific area control case are 0.61 and 0.23, respectively, while those quantities for the western Pacific area control case are 0.79 and 0.28, respectively. Abbreviations used: absorbed solar at the surface (abs.solar), net infrared at the surface (net IR), sensible (sens.), latent and ocean heat advection (adv.); surface moisture convergence (Q-con.), surface moisture (Q), total cloud amount (TCLD), and albedo (alb.).

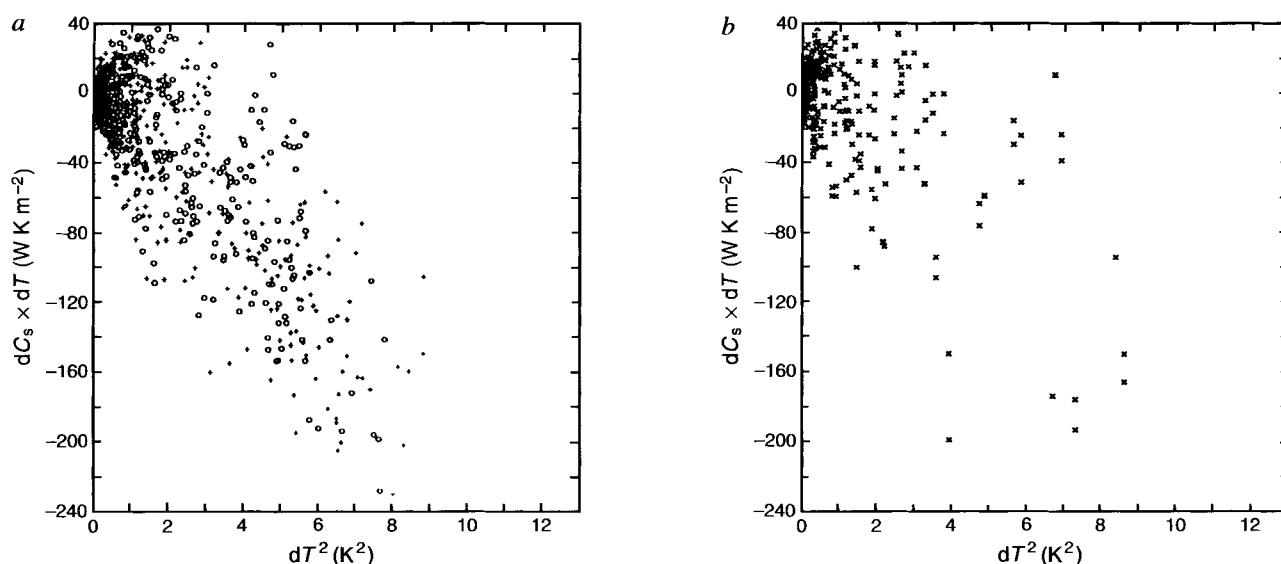


FIG. 1 a, Relative change in solar absorption by the ocean–atmosphere system between warm and cold periods, as measured. Plotted points are the product of the change in SST (dT) and change in cloud radiative forcing at the top of the atmosphere (dC_s as in ref. 8) versus the square of the change in SST (dT^2) for each $2.5^\circ \times 2.5^\circ$ region for a warm (April 1987) minus cold (April 1984) period (+ symbols), and a cold (April 1989) minus warm (April 1987) period (O symbols) in the observations over the tropical Pacific from 10° N to 10° S (ref. 8). Values are plotted for $dC_s \times dT$ less than $+40$ W m⁻² and greater than -240 W m⁻², and dT^2 values between 0

and 13 K². The linear least-squares fit is $dC_s/dT = -19$ W m⁻² K⁻¹ ($r = -0.80$). b, As a but from the the model formulated in exactly the same way except for each $4.5^\circ \times 7.5^\circ$ region (thus there are less samples than in a) for a warm (April, year 63) minus cold (April year 66) period in the equatorial eastern Pacific, and a cold (April year 58) minus warm (April year 63) period in the control run. The linear fit is $dC_s/dT = -13$ W m⁻² K⁻¹ with $r = -0.60$. Both observations and the model show a similar relationship with a negative value for the product indicating a decrease (increase) in absorbed solar energy with a warming (cooling) of the ocean.

threshold can be related physically to static energy considerations such that a surface parcel with SST of 303 K has increased boundary-layer moisture and can be estimated to penetrate through the depth of the tropical troposphere (thus increasing water/ice content and producing brighter middle and high clouds). A surface parcel with, for example, SST of 297 K has less boundary-layer moisture and is only likely to penetrate to the middle troposphere with proportionately less increases of water/ice content in clouds at that level⁸.

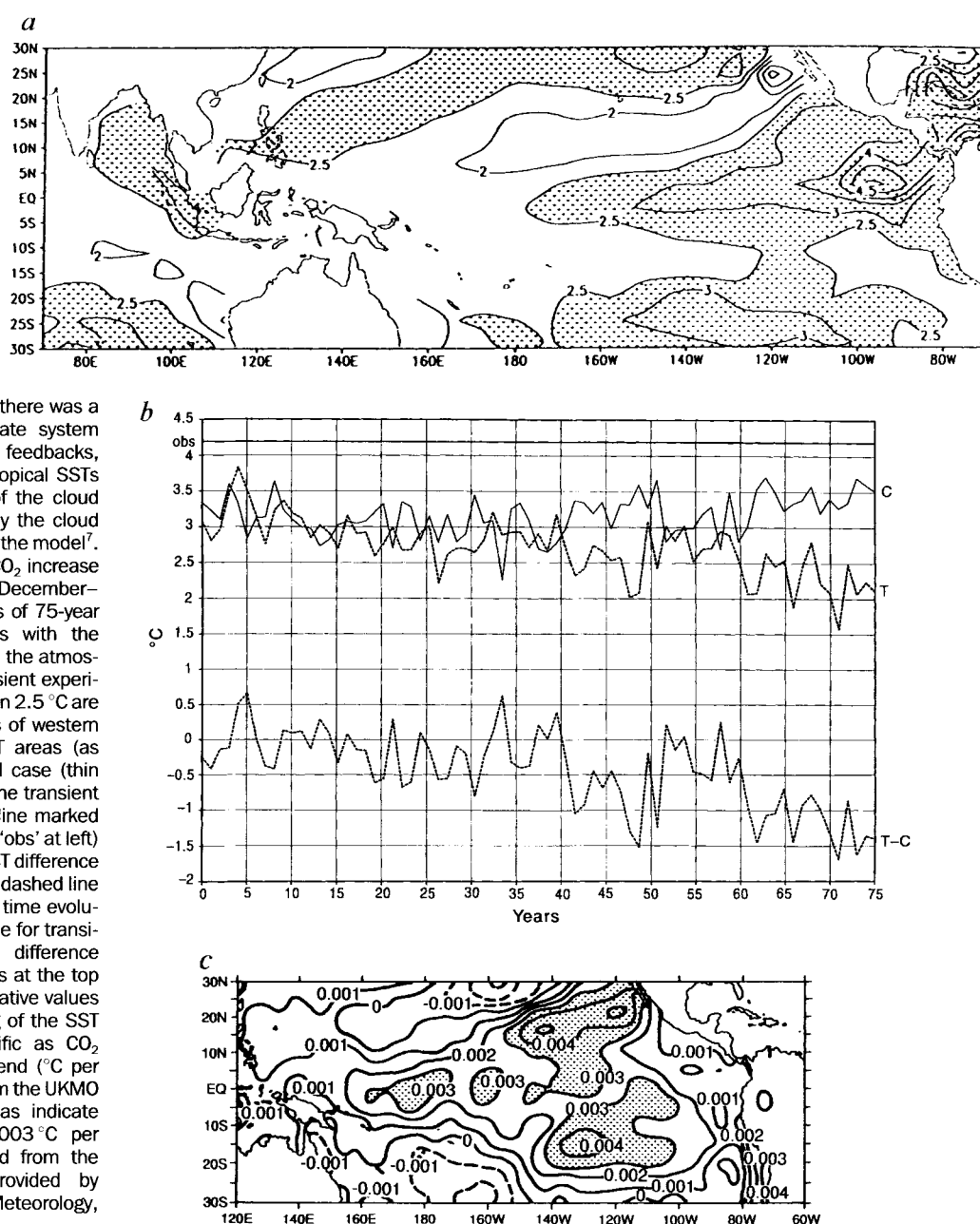
Even though the cloud effects in the present model are simply accounted for, their inclusion compares favourably to a model with a computed cloud optical properties feedback scheme¹⁰ and also reasonably reproduces observed measures of cloud albedo feedback and the super greenhouse effect^{7,8}. A further validation of the simple cloud albedo feedback scheme (Fig. 1) shows that the observed change in cloud shortwave forcing (dC_s) as a function of change in SST (dT) for warm compared to cold episodes in the tropical Pacific (Fig. 1a) produces a linear fit of $-19 \text{ W m}^{-2} \text{ K}^{-1}$

($r = -0.80$). A directly comparable plot for the model (with fewer samples owing to a coarser sampling grid, Fig. 1b) reveals a value of $dC_s/dT = -13 \text{ W m}^{-2} \text{ K}^{-1}$ ($r = -0.60$).

Owing to the uncertainties involved with some cloud processes, cloud feedbacks can be adjusted to produce a wide variety of sensitivities¹¹. But because we do not know all the details of the climate-system response to future changes of forcing, we must assume that parametrizations adjusted to represent present-day conditions (such as the one involving an SST threshold for cloud albedo feedback) would apply to a future climate state. This assumption may not be warranted, and is an important caveat to be kept in mind in this experiment and in any model simulation of climate change with simple parametrizations that represent sometimes controversial interpretations of present-day conditions^{12,13}.

SST anomalies in the tropical Pacific region, increased CO_2 minus control, for the December–January–February season (this season is shown here owing to observational studies that have focused on this season; other seasons in the model show

FIG. 2 Changes in tropical Pacific SSTs from the coupled model and observations. In the coupled model, the atmospheric model has an approximate horizontal resolution of 4.5° latitude and 7.5° longitude with nine vertical levels, the ocean and sea-ice components have 1° latitude–longitude resolution with 20 levels in the ocean, sea ice includes a three-layer thermodynamic scheme along with dynamic sea ice. A sensitivity experiment performed with this version of the coupled model where the cloud albedo feedback was strengthened showed that (1) there was a large-scale response of the climate system combining radiative and dynamic feedbacks, and (2) the maximum values of tropical SSTs were a function of the strength of the cloud albedo feedback as represented by the cloud albedo feedback parametrization in the model⁷. a, SST differences ($^\circ\text{C}$), transient CO_2 increase experiment minus control, for December–January–February, the last 20 years of 75-year transient and control integrations with the coupled model (CO_2 has doubled in the atmosphere at about year 70 of the transient experiment), areas of warming greater than 2.5°C are stippled; b, at top, the time series of western Pacific minus eastern Pacific SST areas (as defined in Table 1) for the control case (thin solid line marked 'C' at right) and the transient CO_2 increase experiment (dashed line marked 'T' at right), thick solid line (marked 'obs' at left) is the observed west-minus-east SST difference as noted in the text; at bottom, the dashed line (marked 'T-C' at right) denotes the time evolution of the west–east SST difference for transient minus control cases (the difference between the solid and dashed lines at the top of the panel), with increasingly negative values indicating a progressive slackening of the SST gradient across the tropical Pacific as CO_2 increases in the model; c, SST trend ($^\circ\text{C}$ per month) for the period 1970–91 from the UKMO GISST dataset^{18,29,30}, stippled areas indicate warming trends greater than 0.003°C per month (SST trend map generated from the GISST dataset cited above, provided by M. Latif, Max Planck Institute for Meteorology, Hamburg, Germany).



similar results) show less warming ($<2^{\circ}\text{C}$, Fig. 2a) where mean SSTs are greatest in the model⁷. Meanwhile in the tropical Pacific east of the Date Line, there is relatively greater warming of the ocean surface ($2\text{--}4^{\circ}\text{C}$). Therefore, the CO_2 -related surface warming is not uniform at the ocean surface across the tropical Pacific. There is a reduction of the east–west or zonal SST gradient along the Equator and immediately to the north, with somewhat less reduction to the south. This slackened equatorial zonal SST gradient is not unlike what is seen during El Niño events^{14–17}. This is not due to a first-order change of frequency of El Niño events in the model, as preliminary results from an analysis of El Niño frequency do not show significant changes between control and increased- CO_2 experiments. A time series of area-averaged differences between eastern and western tropical Pacific SSTs shows a steady decrease of east–west SST gradient throughout the course of the 75-year experiments with increased CO_2 (Fig. 2b).

Observational studies have shown that the SST warming that occurred during the 1980s was not uniform over the tropical Pacific region^{3,17,18}. During this time period when global temperatures also increased, the area-averaged mean warming for the tropical western Pacific ($\sim 0.15^{\circ}\text{C}$ over an area from 120°E to 170°W , 20°N to 20°S) was roughly less than half that in the tropical eastern Pacific ($\sim 0.35^{\circ}\text{C}$ for the area 170°W –coasts of Central and South America, 20°N to 20°S)¹⁸. SST trends from 1970 to 1991 (Fig. 2c) reflect the greater warming east of the Date Line. Elements of this pattern (Fig. 2c) have also been noted in other studies of tropical Pacific SST trends over this period^{13,19,20}. Thus in spite of the some differences in the patterns, both the model and observed trends indicate an overall tropical Pacific warming, with somewhat greater warming east of the Date Line, and a corresponding reduction of zonal SST gradient across the central tropical Pacific.

Associated with this reduction of meridional SST gradient, precipitation increased in the central equatorial Pacific and decreased in the warm-water regime regions of the Intertropical Convergence Zone (ITCZ), South Pacific Convergence Zone (SPCZ), Australasia, and the eastern Indian Ocean in the model

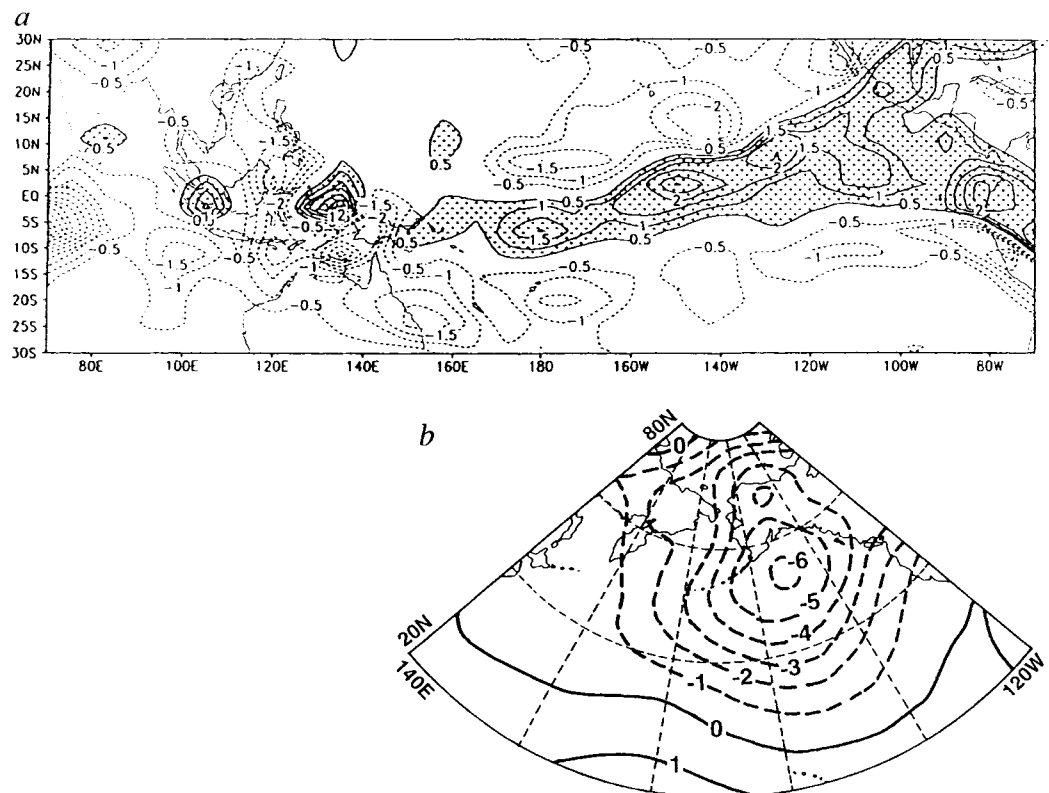
(Fig. 3a). These mean climate-change patterns in the coupled model due to increased CO_2 resemble those associated with represent-day El Niño events in the tropical Pacific region^{15,16}, as well as similar changes observed during the 1980s^{5,17}. Sea-level pressure (SLP) anomaly patterns from the coupled model, increased CO_2 minus control (Fig. 3b), also show the El Niño-like feature seen in the observations of an anomalously deepened Aleutian low-pressure centre in the north Pacific²¹, as well as the similar decadal-timescale SLP anomalies in that region^{2,4,22,23}.

To examine in more detail the mechanisms producing these anomalies in the model, we compiled components of the surface energy balance and low-level moisture parameters for the tropical eastern and western Pacific Ocean areas for 20-year December–January–February averages (Table 1). The changes in SST near the time of CO_2 doubling in the model for these areas correspond to similar changes noted earlier (Fig. 1), with the eastern tropical Pacific area warming more (3.49°C) than the western Pacific (2.21°C).

Absorbed solar radiation at the surface decreases in the western Pacific area (-3.92 W m^{-2}) and increases in the eastern Pacific area ($+2.23\text{ W m}^{-2}$). This is due to cloud albedo feedback effects and changes in cloud amount (see Table 1). Warmer water increases albedo by 1% in the western Pacific area mainly due to cloud albedo feedback, even though total cloud amounts have decreased by 5% in association with the eastward shift of precipitation (Fig. 3). Decreased low clouds and corresponding increases in convective clouds with smaller cloud fraction in the eastern Pacific area contribute a decrease of total albedo by 4% (total cloud amount decreases by 13%, due to a decrease of low-level, predominantly stratus, clouds of 27%; middle- and upper-level, predominantly cumulus, clouds increase by 123%).

There are small changes in net infrared radiation and sensible (due to the temperature difference between the surface and overlying air) heat flux, but there is a large increase of latent heat flux (positive sign denotes heat removed from the ocean surface) in the eastern Pacific area compared to the western Pacific area ($+8.67\text{ W m}^{-2}$ versus $+0.81\text{ W m}^{-2}$). This is due to

FIG. 3 a, Precipitation differences (mm d^{-1}), transient CO_2 increase experiment minus control, for December–January–February, the last 20 years of 75-year transient and control integrations with the coupled model (CO_2 has doubled in the atmosphere at about year 70 of the transient experiment), areas of precipitation increase are grey-shaded, and this pattern is similar to the precipitation anomalies during typical El Niño events³²; note other coupled-model study of anomalously dry areas in the tropics during El Niño events becoming drier with increased CO_2 (ref. 34). b, Sea level pressure differences (mbar) for the north Pacific region, transient CO_2 increase experiment minus control, for December–January–February, the last 20 years of 75-year transient and control integrations with the coupled model (CO_2 has doubled in the atmosphere at about year 70 of the transient experiment).



greater evaporation in the eastern Pacific associated with larger SST increases, as well as relatively large increases of surface moisture (+16%) in the western Pacific that contribute to inhibiting larger evaporation increases there. There is also a large increase in low-level moisture convergence in the eastern Pacific area (+62%) compared to the western area (+3%). But, in absolute terms, there is still somewhat greater latent heat flux and low-level moisture amounts in the western Pacific area compared to the eastern area as observed^{24,25}, due to the higher base-state SSTs in the west⁷. Thus the changes in latent heat flux are the result of factors that produce an effect opposite to what would contribute to a reduction of zonal SST gradient.

Changes in ocean advection (mostly due to changes in upwelling; weak horizontal SST gradients in these particular regions do not contribute significantly to horizontal heat advection⁷) are larger for the eastern compared to the western Pacific area, indicating that the weakened easterlies and associated decreased upwelling (not shown) result in greater heating in the eastern Pacific (+7.56 W m⁻² compared to +5.28 W m⁻²) where upwelling is stronger in the present-day climate case.

Thus radiative changes (partly due to cloud albedo feedback, and partly due to changes in cloud amount), plus the large-scale dynamical atmosphere-ocean feedbacks that reduce the surface easterly winds and weaken equatorial upwelling in the ocean, combine to contribute to the greater relative warming of tropical Pacific SSTs east of the Date Line, with attendant precipitation and circulation anomalies, due to increased CO₂ in the coupled model. Droughts that already occur in the Australasia/western Pacific region in association with El Niño events could thus intensify, disrupting water resources on small islands dependent on rainfall for fresh water, and contributing to long-term depletion of those resources⁶. We cannot definitively attribute the recent warming in the Pacific (and associated global warming of the 1980s and early 1990s) to increased levels of CO₂ in the atmosphere. Yet the model results presented here and elsewhere^{26,27}, as well as some observational results²⁸, point to the possibility that CO₂-induced climate change in the Pacific region could have this signature. However, there may also be decadal-timescale variability in the tropical Pacific region that could complicate the interpretation of this signal (M. Latif, personal communication). Further clarification of these effects awaits more definitive observational and modelling studies, improved cloud formulations, and a better understanding of observed decadal-timescale climate fluctuations in the Pacific region. □

Received 12 June 1995; accepted 21 May 1996.

- Nitta, T. & Yamada, S. *J. met. Soc. Japan* **67**, 375–383 (1989).
- Trenberth, K. E. & Hurrell, J. W. *J. Clim. Dyn.* **9**, 303–319 (1994).
- Graham, N. E. *Clim. Dyn.* **10**, 135–162 (1994).
- Nitta, T. & Kachi, M. *J. met. Soc. Japan* **72**, 823–831 (1994).
- Kumar, A., Leetma, A. & Ji, M. *Science* **266**, 632–634 (1994).
- Meehl, G. A. *J. Wat. Air Soil Pollut.* (in the press).
- Meehl, G. A. & Washington, W. M. *Clim. Dyn.* **11**, 399–411 (1995).
- Ramanathan, V. & Collins, W. *Nature* **351**, 27–32 (1991).
- Washington, W. M. & Meehl, G. A. *Clim. Dyn.* **8**, 211–223 (1993).
- Boer, G. *J. Clim. Dyn.* **8**, 225–239 (1993).
- Senior, C. A. & Mitchell, J. F. B. *J. Clim.* **6**, 393–418 (1993).
- Houghton, J. T., Callander, B. A. & Varney, S. K. (eds) *Climate Change 1992: The IPCC Scientific Assembly Supplementary Report* (Cambridge Univ. Press, 1992).
- Wallace, J. M. *Nature* **357**, 230–231 (1992).
- Rasmusson, E. M. & Carpenter, T. H. *Mon. Weath. Rev.* **110**, 354–384 (1982).
- Kiladis, G. N. & van Loon, H. *Mon. Weath. Rev.* **116**, 120–136 (1988).
- Meehl, G. A. *Mon. Weath. Rev.* **115**, 27–50 (1987).
- Salinger, J. M. et al. *Int. J. Climatology* **15**, 285–302 (1995).
- Bottomly, M., Folland, C. K., Hsiung, J., Newell, R. E. & Parker, D. E. *Global Ocean Surface Temperature Atlas* (UK Met. Office, Bracknell, 1990).
- Kito, A. *J. met. Soc. Japan* **73**, 247–253 (1995).
- Gutzler, D. S. *J. Atmos. Sci.* (in the press).
- van Loon, H. & Madden, R. A. *Mon. Weath. Rev.* **109**, 1150–1162 (1981).
- Trenberth, K. E. *Bull. Am. met. Soc.* **71**, 988–993 (1990).
- Chen, T.-C., van Loon, H., Wu, D.-D. & Yen, M.-C. *J. met. Soc. Japan* **70**, 1137–1146 (1992).
- Oberhuber, J. M. *An atlas Based on the COADS Data Set: The Budgets of Heat, Buoyancy and Turbulent Kinetic Energy at the Surface of the Global Ocean* (Max Planck Institute for Meteorology, Hamburg, 1988).
- Zhang, G. J. & McPhaden, M. J. *J. Clim.* **8**, 589–605 (1995).
- Knutson, T. R. & Manabe, S. *J. Clim.* **8**, 2181–2199 (1995).
- Tett, S. *J. Clim.* **8**, 1473–1502 (1995).
- Trenberth, K. E. & Hoar, T. *J. Geophys. Res. Lett.* **23**, 57–60 (1996).
- Folland, C. K., Parker, D. E. & Kates, F. E. *Nature* **310**, 670–673 (1984).

- Folland, C. K. & Parker, D. E. *Climate–Ocean Interaction* 21–52 (Kluwer, Dordrecht, 1990).
- Meehl, G. A., Branstator, G. W. & Washington, W. M. *J. Clim.* **6**, 42–63 (1993).
- Ropelewski, C. F. & Halpert, M. S. *Mon. Weath. Rev.* **115**, 1606–1626 (1987).
- Alexander, R. C. & Mobley, R. L. *Mon. Weath. Rev.* **104**, 143–148 (1976).
- Meehl, G. A. & Washington, W. M. *Mon. Weath. Rev.* **114**, 667–674 (1986).
- Graham, N. *Science* **267**, 666–671 (1995).

ACKNOWLEDGEMENTS. We thank M. Latif for comments and for providing the observed SST trend map, and J. Hurrell for his comments. Some of the computations were performed under the auspices of the Model Evaluation Consortium for Climate Assessment (MECCA) and CRAY Research, Inc. The National Center for Atmospheric Research is sponsored by the National Science Foundation. This work was supported in part by the US Department of Energy, Office of Health and Environmental Research as part of its Carbon Dioxide Research Program.

CORRESPONDENCE should be addressed to G.A.M. (e-mail meehl@ncar.ucar.edu).

Life and times of the Bering land bridge

Scott A. Elias*, Susan K. Short*, C. Hans Nelson† & Hilary H. Birks‡

* Institute of Arctic and Alpine Research, Campus Box 450, University of Colorado, Boulder, Colorado 80309-0450, USA

† US Geological Survey, Mail Stop 999, Menlo Park, California 94025, USA

‡ Botanical Institute, University of Bergen, Allegaten 41, N-5007 Bergen, Norway

UNDERSTANDING the environment of the Bering land bridge and determining the timing of late Wisconsin inundation are important for several areas of study. These include: (1) the timing of the re-establishment of circulation between Pacific and Atlantic Oceans; (2) the timing of development of a northern biotic refugium and the closing of the bridge to species immigration; (3) Palaeoindian migration routes; and (4) palaeotopographic data for atmospheric general circulation models¹. Late Wisconsin palaeobotanical and fossil insect data from the central and northern sectors of the Bering land bridge indicate widespread mesic shrub tundra environments even during the last glacial maximum. Contrary to previous hypotheses, we found no evidence of steppe tundra on the land bridge. New accelerator mass spectrometer ¹⁴C dates show much of the land bridge was above sea level and thus available for human and animal migration until 11,000 yr BP. Insect evidence suggests that summer temperatures at that time were substantially warmer than now.

The reconstruction of full-glacial vegetation of Beringia has been controversial^{2–5}. The hypothesis of a productive grassland was based on fossils of large grazers in full-glacial sediments. High percentages of *Artemisia* in Beringian pollen spectra were interpreted as being attributable to a steppe tundra or mammoth steppe, dominated by plants found in Asian steppes today⁴. The steppe-tundra hypothesis has been questioned on the basis of palynology^{2,3,5}. Prior reconstructions of lowland environments on the land bridge have been based on extrapolation from adjoining upland regions, so no definitive evidence for the full-glacial vegetation of the land bridge has been available.

We analysed 20 cores collected by the US Geological Survey from the Bering and Chukchi Seas for beetles, pollen, and plant macrofossils to reconstruct environments during the last glaciation (Fig. 1). The organic strata, ranging in thickness from <5 to 300 cm, were *in situ* terrestrial peats and organic detrital silts formed on the emergent land bridge, subsequently overlain by Holocene marine sediments from marine transgression.

Radiocarbon dates from these cores range in age from 3,455 to >49,200 yr BP (Table 1). New accelerator mass spectrometer (AMS) ages from peat near the top of non-marine sediment on the Chukchi shelf suggest that the previous conventional radiocarbon ages on bulk sediments were as much as 4.25 kyr too old⁶.

McManus and Creager⁷ estimated that the land bridge was inundated by 14,400 ¹⁴C yr BP. However, their sea-level chronology is suspect because it was based on radiocarbon ages of bulk

TABLE 1 Radiocarbon date list, core and sample data from the Bering and Chukchi shelves

Group	Core	Site label (Fig. 1)	Location	Water depth (m)	Depth in core (cm)	¹⁴ C age (yr BP)	Material dated	Laboratory number
Port Clarence	76-71A	i	65°05' N 167°44' W	34	70–75	13,200 ± 110	Peat	USGS-155
	76-71A	i		34	105–125	14,035 ± 105*	Screened peat	AA-14671
	76-72	f	65°08' N 167°36' W	34	40	15,450 ± 250	Peaty sediment	USGS-537
	76-101	g	65°07' N 167°31' W	28	80–86	16,400 ± 450	Peat	USGS-157
	76-101	g		28	120	16,540 ± 200	Peaty sediment	USGS-356
	76-101	g		28	80–120	12,350 ± 85*	Screened peat	AA-14672
	78-15	e	65°14' N 167°25' W	26	135	13,200 ± 110	Peaty sediment	USGS-761
	78-15	e		26	160	12,600 ± 100*	Peaty sediment	USGS-760
	78-15	e		26	238	13,250 ± 85*	Peaty sediment	USGS-727
	78-15	e		26	225–233	11,030 ± 85*	Screened peat	AA-14676
	78-15	e		26	415–420	20,725 ± 165*	Screened peat	AA-14675
	78-17	h	65°13' N 167°28' W	17	185	46,100 ± 1,800	Peaty sediment	USGS-762
Mid-Chirikov Basin	77-37	j	64°34' N 169°23' W	40	142–146	40,350 ± 1,300*	Screened peat	AA-14673
	80-33	k	64°16' N 168°42' W	41	70–83	13,715 ± 130*	Screened peat	AA-14678
Nome	76-121	q	63°53' N 163°02' W	17	80–86	11,570 ± 130	Peaty sediment	USGS-358
	76-121	q		17	175–180	29,500 ± 340	Peaty sediment	USGS-158
	76-131	o	64°24' N 161°50' W	17	130–135	11,800 ± 200	Peaty sediment	USGS-159
	77-15	p	64°10' N 165°28' W	20	134–140	13,770 ± 210	Peat	USGS-352
	78-3	l	64°06' N 165°30' W	18	343	10,280 ± 50	Peaty sediment	USGS-758
	78-3	l		18	330	11,080 ± 60*	Screened peat	NSRL-2398
	78-4	m	64°11' N 165°33' W	18	112	12,580 ± 90	Peat	USGS-857
	78-6	n	64°07' N 164°20' W	20	215	32,000 ± 800	Peaty sediment	USGS-725
	78-6	n		20	150	>49,200*	Screened peat	AA-14674
Yukon Delta	78-22	t	63°21' N 165°50' W	14	344	13,870 ± 85	Peaty sediment	USGS-728
	78-23	u	63°13' N 165°12' W	24?	125	4,700 ± 30	Peaty sediment	USGS-729
	80-3	s	63°29' N 164°53' W	11	90–92.5			
	80-3	s			102–108			
	80-4	r	63°31' N 164°56' W	8	45–50	3,455 ± 55*	Screened peat	AA-14677
	80-4			8	45–50	4,820 ± 60*	Insect fossils	NSRL-2399
Chukchi Sea	85-54	b	71°03' N 166°15' W	46	450–460	11,330 ± 70*	Screened peat	Beta-43952
	85-58	a	71°06' N 167°36' W	48	185–190	12,215 ± 150*	Screened peat	AA-14679
	85-69	c	70° N 165°39' W	44	90–95	11,000 ± 60*	Screened peat	Beta-43953
	85-70	d	69°57' N 165°21' W	42	220–230	13,195 ± 90*	Screened peat	AA-14680

* AMS dates.

sediment containing 'coal-like material'. We wet-screened samples to remove such contaminants, and obtained AMS ¹⁴C ages on large macrofossils. This method provided a more accurate chronology because true ages for land bridge peats are several thousand years younger than ¹⁴C ages based on unwashed bulk samples (for example, see Table 1, core 76-101 or 233–238 cm in core 78-15).

In addition to old carbon contamination⁷ problems with Bering sea-level reconstruction (Fig. 2), wave erosion may also have a greater effect there. In the Chukchi Sea, most samples came from water depths >30 m, whereas most Bering Sea samples were from <30 m water depth (Fig. 1). Wave erosion of peat after sea-level transgression was probably greater in the Bering Sea because storms often scour the shallower sea bottom there⁸. Thus, the highest level of the pre-transgressive peat dated in Bering Sea is probably relatively older than that in the Chukchi Sea, resulting in an earlier reconstructed rise of sea level over the Bering shelf (Fig. 2). On the basis of these arguments, the best ¹⁴C dates to determine land-bridge inundation history derive from the Chukchi shelf.

The youngest AMS radiocarbon ages found at the greatest water depths on the Chukchi shelf are at sites b and c (Figs 1, 2). These data do not define a complete sea-level curve, but they provide the most accurate information compared to global

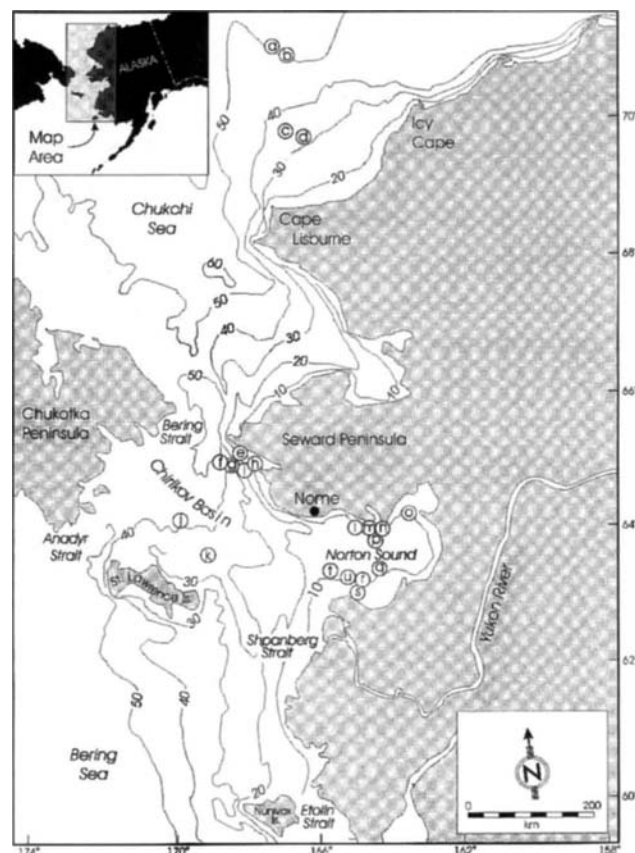


FIG. 1 Map of the northwest coast of Alaska, showing core localities on the Chukchi (a–d) and Bering sea (e–u) shelves. Radiocarbon dates from the uppermost terrestrial peats in Chukchi shelf cores 54 and 69 were used to establish the timing of inundation of the land bridge.

eustatic sea level curves⁹⁻¹² (Fig. 2) at the critical pre-Holocene sill depths of about 48 m in the Anadyr^{8,13} and Herald Straits⁷. These straits control the opening and closing of the land bridge.

Relatively older AMS peat ages from Bering shelf may suggest temporary transgressions through the Anadyr Strait before formation of a permanent continuous seaway through the land bridge (Table 1; Fig. 2). Such transgressions may have extended through the deeper Bering Strait to form an embayment into the southern Chukchi Sea, where pre-Holocene palaeo-bathymetry is 50–60 m¹⁴ (Fig. 1). Chukchi sites b and c, which bracket the northern –48 m sill depth, however, indicate that the permanent seaway from the Pacific to the Arctic Ocean was not established until 11 kyr BP.

Based on analysis of our Chukchi samples, sea level rose to –40 m about 10.5 kyr BP (Fig. 2). At this time, Anadyr Strait was about 30 km wide, Bering Strait approached its present width, and most of the Chukchi Sea was flooded. This has important palaeo-climatic implications because warm Pacific water was free to circulate into the Arctic Ocean, covering nearly 400,000 km² of the former land bridge. Biotic migration routes were only restricted by the narrower western strait waterways at this time, and would not have been eliminated until shelf inundation reached –30 m depth about 9.5 kyr BP. At this time, the Shpanberg Strait, east of St Lawrence island, had a width equal to the present Bering Strait. By 5 kyr BP (Fig. 2, site r), the northern Bering shelf flooded, nearly completing land-bridge inundation.

The fossil samples fall into three age classes: more than 40,000 yr BP, 20,000–14,000 yr BP and 14,000–11,000 yr BP (Table 2). The majority of 64 pollen samples date to the Alaskan Birch

interval (14,000–9,000 yr BP)^{15,16}; three cores record full glacial environments (Table 2). All pollen spectra are characterized by large *Betula* (10–60%), Cyperaceae (5–41%), Poaceae (20–46%) and *Sphagnum* (5–58%) percentages. *Betula* percentages increased from mid-Wisconsin through late-glacial times. There is little or no *Alnus*, *Picea* or *Artemisia* pollen. Mesic shrub graminoid tundra was widespread across the land bridge in all time periods; birch was important especially in the late glacial, and heaths and shrub willows were locally abundant. The central and northern sectors were characterized by fewer birch shrubs and more abundant grasses, and pollen concentration values decrease with increasing distance from the present coast, suggesting greater continentality in the centre of the land bridge. Pond and coastal marsh environments are indicated by *Typha*/*Sparganium*, *Plantago maritima* and *Triglochin* pollen.

The plant macrofossils (Table 2) suggest a landscape mosaic. Shallow-water taxa include *Ranunculus*, *Potamogeton vaginatus* and *Myriophyllum spicatum*. Emergent marginal communities contained *Hippuris vulgaris*, *Menyanthes trifoliata* and *Carex* spp., probably associated with the brown mosses *Scorpidium scorpioides* and *Drepanocladus*. At or above the water table, wet *Sphagnum*-dominated communities supported *Andromeda polifolia*. Drier upland communities are suggested by the dwarf shrub *Empetrum* cf. *hemaphroditum* and *Silene*. Few macrofossils of upland taxa were recorded, suggesting that there was little water inflow or possibly early and long-lasting snow cover, preventing dispersal of upland seeds. No steppe or grassland taxa were found. Recent Alaskan studies suggest that the steppe-tundra biome was not as widespread as first postulated^{17,18}. Steppe tundra probably did not

TABLE 2 Summary of palaeoenvironments and fossil data from the central and northern sectors of the Bering land bridge from 40,000–11,000 yr BP, divided into three time intervals

Time span	Pollen spectra	Plant macrofossils	Insect fauna	Palaeoenvironment
>40,000 yr BP	15–20% <i>Betula</i> 20–46% Poaceae 5–13% Cyperaceae 5–20% <i>Sphagnum</i> 2–3% <i>Salix</i> 1–5% Ericaceae	Samples too small for macrofossil study	Samples too small for insect fossil study	Vegetation a birch–heath–graminoid tundra with little or no steppe elements. Shrubs not an important element of vegetation, but present in small numbers
20,000–14,000 yr BP	19–28% <i>Betula</i> 28–44% Poaceae 13–30% Cyperaceae 1–5% <i>Salix</i> 5–13% Ericaceae 26–58% <i>Sphagnum</i>	<i>Scorpidium scorpioides</i> <i>Drepanocladus</i> sp. <i>Sphagnum</i> spp. <i>Hippuris vulgaris</i> <i>Carex tristigmatae</i> <i>Empetrum nigrum</i> <i>Menyanthes trifoliata</i> <i>Andromeda polyfolia</i> <i>Myriophyllum spicatum</i> <i>Potamogeton vaginatus</i>	<i>Pterostichus nivalis</i> <i>Colymbetes dolabratus</i> <i>Olophrum</i> sp. <i>Tachinus brevipennis</i>	Vegetation was dominated by birch–graminoid tundra with small ponds choked with aquatic plants. Heaths relatively unimportant. No evidence for steppe-tundra vegetation. Insects indicative of arctic climate, with drier tundra than during the late glacial. No steppe-tundra insect species
14,000–11,000 yr BP	15–60% <i>Betula</i> 20–40% Poaceae 10–30% Cyperaceae <24% Ericaceae 1–5% <i>Salix</i> 5–50% <i>Sphagnum</i>	<i>Scorpidium scorpioides</i> <i>Drepanocladus</i> sp. <i>Betula nana</i> <i>Silene</i> sp. <i>Hippuris vulgaris</i> <i>Carex tristigmatae</i> <i>Empetrum nigrum</i> <i>Andromeda polyfolia</i> <i>Ledum decumbens</i> <i>Menyanthes trifoliata</i> <i>Myriophyllum spicatum</i> <i>Potamogeton berchtoldii</i> <i>P. vaginatus</i> <i>Ranunculus</i> cf. <i>trichophyllum</i> <i>Ranunculus</i> cf. <i>hyperboreus</i> <i>Potentilla palustris</i> <i>Ribes</i> sp.	<i>Dyschirius nigricornis</i> <i>Bembidion grapsei</i> <i>Diacheila polita</i> <i>Pterostichus agonus</i> <i>P. brevicornis</i> <i>P. caribou</i> <i>P. haematopus</i> <i>P. sublaevis</i> <i>Colymbetes dolabratus</i> <i>Agabus moestus</i> <i>Hydroporus polaris</i> <i>Helophorus splendidus</i> <i>Holoboreaphilus nordenskiöldi</i> <i>Tachinus brevipennis</i> <i>Tachyporus rufomus</i> <i>Pycnoglypta lurida</i> <i>Gymnusa atra</i> <i>Simplocaria tessellata</i> <i>Lepidophorus lineaticollis</i> <i>Isochnus flagellum</i>	Vegetation was dominated by birch–heath–graminoid tundra with small ponds choked with aquatic plants. Insects indicate open ground habitats dominated by mesic tundra. By 12,000 yr BP, summer temperatures were as warm as present; by 11,000 yr BP, summer temperatures were warmer than present-day north slope of Alaska

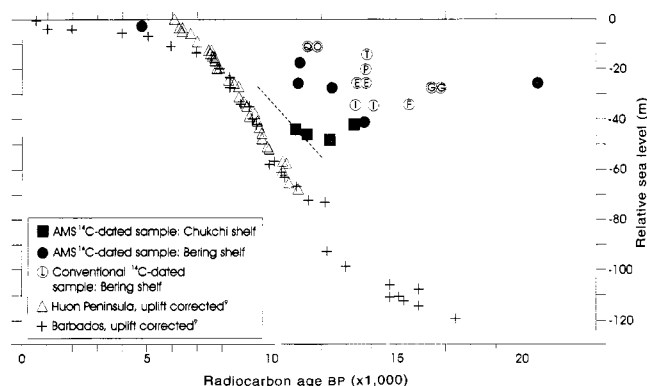


FIG. 2 Sea-level curve comparison (see Table 1 and Fig. 1 for core identifications). Note that in key Chukchi Sea core sites 54 (11.3 kyr) at 50.7 m and 69 (11 kyr) at 45.5 m, the plant macrofossils were collected from peat layers that were 20 (core 69) to 60 cm (core 54) below the apparent change from non-marine to marine sediment. Based on the peat sedimentation rates ($20\text{--}100\text{ cm yr}^{-1}$) in other cores (for example, 78–15 in Table 1), the ages of the youngest emergent peat and marine inundation at sites 54 and 69 may be a minimum of several hundred years younger. All ages shown are in radiocarbon years BP.

dominate the land bridge; mesic tundra was far more widespread than previously thought.

The insect assemblages, including many *Pterostichus* (*Cryobius*) species of ground beetles and omaliine rove beetles, represent mesic tundra environments (Table 2). The aquatic and semi-aquatic fauna represent sedge- and sphagnum-lined ponds. With the exception of *Lepidophorus lineaticollis*, no steppe-tundra taxa¹⁹ were recovered.

Beetles dating to 20,000 yr BP are indicative of relatively dry tundra heaths or dry meadows. The full glacial insect data indicate cold, moderately dry to mesic environments, with heaths, dry meadows and shrub tundra, interspersed with marshes and small ponds.

Late glacial insect assemblages (14,000–11,000 yr BP) reflect increasingly mesic and moist habitats and warming climates just before inundation (Table 2). Late glacial assemblages reflect climatic warming from full glacial conditions to warmer-than-modern summers. We identified boreal taxa rarely found today beyond the treeline^{20–24}. Mutual climatic range (MCR) analysis of Bering shelf assemblages from 11,500–11,000 yr BP yielded mean summer temperatures (T_{max}) ranging from 11.5 to 13 °C and mean winter temperatures (T_{min}) of –28.5 to –23.5 °C. Mean July temperatures at Wales and Nome are 8.3 and 10.8 °C, respectively; mean February temperatures are –20.3 and –15.6 °C, respectively²⁵. Thus the late glacial climate on the Bering shelf was characterized by summer temperatures that were somewhat warmer than modern mainland conditions, but winter temperatures remained considerably colder than modern.

Chukchi shelf insect assemblages from localities >500 km north of the Bering shelf from 11,500–11,000 yr BP yielded MCR results that reflected slightly cooler conditions. MCR analysis of the late glacial Chukchi core assemblages yielded T_{max} values in the range 9.25 to 10.25 °C and T_{min} values from –32.25 to –27.25 °C. At Barrow, mean July temperature is 4 °C and mean February temperature is –27.7 °C (ref. 25). Late glacial summers on the Chukchi shelf were therefore considerably warmer than modern (onshore) parameters, and late glacial winters were approaching modern parameters. This 'warm' signal corresponds to the time of maximum postglacial insolation at 60° N (ref. 26) and was also recorded from contemporaneous and early Holocene insect assemblages in arctic Alaska and the Yukon territory^{27–29}. There is no botanical evidence for warmer-than-modern climates during this interval. New radiocarbon dates from the land bridge indicate that its inundation coincided with this warm interval, rather than

predating it by several thousand years, as suggested by McManus and Creager's sea-level curve⁷. □

Received 12 March; accepted 14 May 1996.

1. Peltier, W. R. *Science* **265**, 195–201 (1994).
2. Colinvaux, P. *Quat. Res.* **16**, 18–36 (1981).
3. Cwynar, L. C. & Ritchie, J. C. *Science* **208**, 1375–1377 (1980).
4. Guthrie, R. D. *Frozen Fauna of the Mammoth Steppe* (University of Chicago Press, Chicago, Illinois, 1990).
5. Ritchie, J. C. *Past and Present Vegetation of the Far Northwest of Canada* (University of Toronto Press, Toronto, 1984).
6. Elias, S. A., Short, S. K. & Phillips, R. L. *Quat. Res.* **38**, 371–378 (1992).
7. McManus, D. A. & Creager, J. S. *Quat. Res.* **21**, 317–325 (1984).
8. Nelson, C. H. *Geol. Mijnbouw* **61**, 5–18 (1982).
9. Chapell, J. & Polach, H. *Nature* **349**, 147–149 (1991).
10. Edwards, R. L. et al. *Science* **260**, 962–968 (1993).
11. Mann, D. H. & Peteet, D. M. *Quat. Res.* **42**, 136–148 (1994).
12. Fairbanks, R. G. *Nature* **342**, 637–647 (1989).
13. Hopkins, D. M., Nelson, C. H., Perry, R. B. & Alpha, T. R. *US Geol. Surv. Prof. Paper* **759-B** (1976).
14. Holmes, M. L. thesis, Univ. of Washington (1975).
15. Livingstone, D. A. *Ecology* **36**, 587–600 (1955).
16. Anderson, P. M., Bartlein, P. J. & Brubaker, L. *Quat. Res.* **41**, 306–315 (1994).
17. Lea, P. D., Elias, S. A. & Short, S. K. *Arct. Alp. Res.* **23**, 375–391 (1991).
18. Elias, S. A. *Arct. Alp. Res.* **24**, 133–144 (1992).
19. Matthews, J. V. in *Paleoecology of Beringia* (eds Hopkins, D. M., Matthews, J. V. Jr, Schweger, C. E. & Young, S. B.) 127–150 (Academic, New York, 1982).
20. Lindroth, C. H. *Opusc. Entomol. Suppl.* **24**, 201–408 (1963).
21. Lindroth, C. H. *Opusc. Entomol. Suppl.* **20**, 1–200 (1961).
22. Campbell, J. M. *Can. Entomol.* **116**, 487–527 (1984).
23. Campbell, J. M. *Can. Entomol.* **115**, 361–370 (1983).
24. Campbell, J. M. *Mem. Entomol. Soc. Can.* **109**, 1–95 (1979).
25. National Oceanic and Atmospheric Administration *Climatology of the United States. Monthly Normals of Temperature, Precipitation, and Heat and Cooling Degree Days by State, 1951–1980* (Washington DC, 1982).
26. Bartlein, P. J., Anderson, P. M., Edwards, M. E. & McDowell, P. F. *Quat. Internat.* **10–12**, 73–83 (1991).
27. Matthews, J. V. Jr *Geol. Surv. Can. Pap.* **75-1B**, 139–146 (1975).
28. Wilson, M. J. & Elias, S. A. *Arctic* **39**, 150–157 (1986).
29. Nelson, R. E. & Carter, L. D. *Arct. Alp. Res.* **19**, 230–241 (1987).

ACKNOWLEDGEMENTS. We thank L. Phillips, US Geological Survey, Menlo Park, California, for access to cores from the Chukchi shelf and for additional core data, and D. Hopkins, University of Alaska, and B. Richmond, US Geological Survey, Menlo Park, for reviewing a first draft of the manuscript. This research was funded by the Office of Polar Programs, NSF. MCR reconstructions were funded by the Climate Dynamics Program, NSF.

CORRESPONDENCE and requests for materials should be addressed to S.A.E. (e-mail: saelias@cultur.colorado.edu).

Insect motion detectors matched to visual ecology

D. C. O'Carroll, N. J. Bidwell*, S. B. Laughlin & E. J. Warrant†

Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK

* Centre for Neuroscience, University of Sussex, Falmer, Brighton BN1 9QG, UK

† Department of Zoology, University of Lund, Helgonavägen 3, S-22362 Lund, Sweden

To detect motion, primates, birds and insects all use local detectors to correlate signals sampled at one location in the image with those sampled after a delay at adjacent locations^{1–10}. These detectors can adapt to high image velocities by shortening the delay^{11–13}. To investigate whether they use long delays for detecting low velocities, we compared motion-sensitive neurons in ten species of fast-flying insects, some of which encounter low velocities while hovering. Neurons of bee-flies and hawkmoths, which hover, are tuned to lower temporal frequencies than those of butterflies and bumblebees, which do not. Tuning to low frequencies indicates longer delays and extends sensitivity to lower velocities. Hoverflies retain fast temporal tuning but use their high spatial acuity for sensing low-velocity motion. Thus an unexpectedly wide range of spatio-temporal tuning matches motion detection to visual ecology.

In the blowfly *Calliphora*, wide-field, direction-selective neurons of the third optic ganglion are a major component of the optomotor pathway for flight stabilization¹⁴. We used drifting, sinusoidal gratings to determine the temporal and spatial frequency tuning of similar neurons in ten insect species from

three orders. All species are capable of fast forward flight, but some are noted for their ability to hover. All neurons recorded had a strongly direction-selective response to wide-field motion (Fig. 1) and no preference for small targets.

How do such neurons detect the low image velocities encoun-

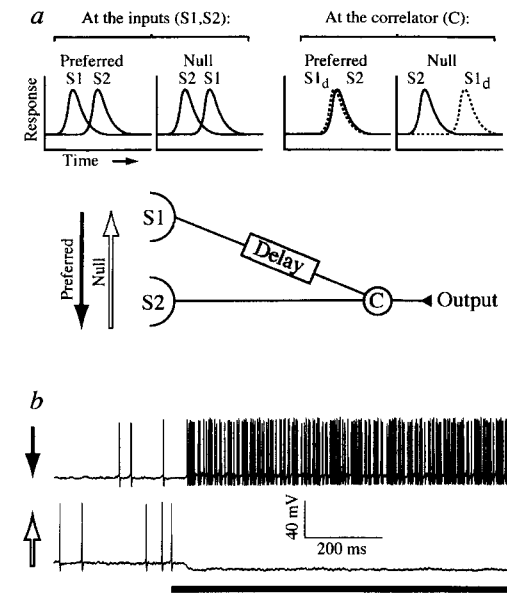


FIG. 1 a, The local correlation scheme used to account for the direction-selective responses of neurons to image motion¹⁻¹⁰. Two sampling stations (S1 and S2) view adjacent points in space and project to a correlator (C). An object moving past these inputs generates sequential responses, either first at S1 (downward motion) or first at S2 (upward motion). The arrival of the response from S1 at the correlator (C) is delayed by a mechanism such as a low-pass filter¹⁻³. If the object moves from S1 to S2 in a time equal to the delay, the delayed and undelayed inputs, S1_d and S2, coincide at the correlator (C) (top right). When the object moves in the opposite (null) direction, however, the delay mechanism increases the separation between the two inputs. Correlation involves a nonlinear operation resembling multiplication, so the output signal is greatly enhanced by coincidence¹⁻³. b, Responses recorded intracellularly from a neuron in the lobula plate of a syrphid fly, *Eristalis tenax*. The physiology of these neurons is consistent with excitatory input from an array of correlators, as in a. This neuron is sensitive to downward motion, giving strong responses when a pattern is moved in the preferred direction (top; stimulus duration indicated by dark bar), but is inhibited by motion in the opposite direction (bottom).

METHODS. All experiments were performed by recording intracellularly from neurons in the lobula complex of animals restrained by wax, using aluminium silicate glass micropipettes filled with 1 M KCl (tip resistance $\sim 100 \text{ M}\Omega$). Stimuli were presented on an XYZ monitor (Tektronix 608) at a frame rate of 200 Hz, driven by a Picasso Image Synthesiser, under computer control. The screen was positioned 7 cm in front of the insect, and centred on the receptive field of the neuron, allowing the production of patterns subtending approximately 70° .

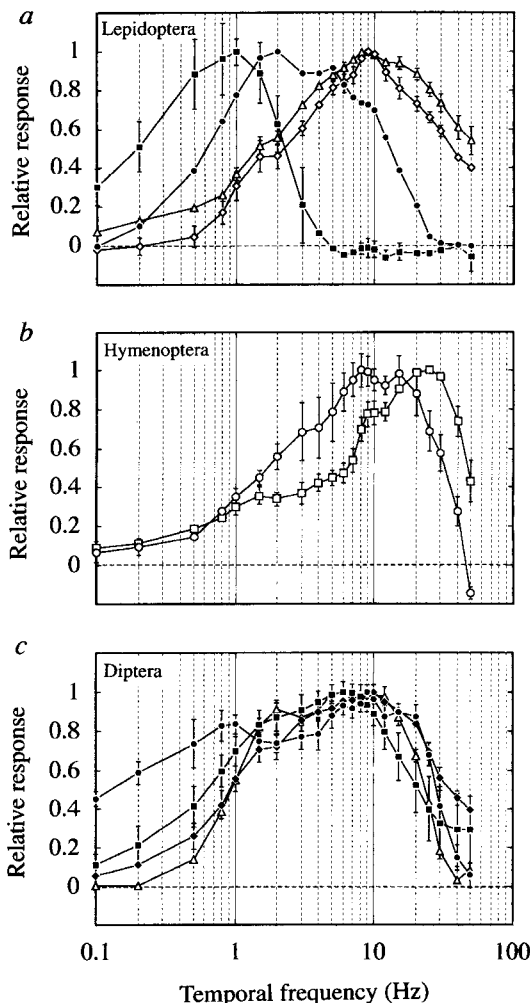


FIG. 2 Temporal frequency tuning curves determined for neurons in an unadapted state, as encountered when a stationary animal starts moving. Data are from ten species from three orders: a, Lepidoptera; b, Hymenoptera; c, Diptera. a, *Hemaris fuciformis* ($n = 1$, filled circles), hawkmoth *Deilephila elpenor* ($n = 3$, filled squares), *Inachis io* ($n = 3$, open diamonds) and butterfly *Vanessa atalanta* ($n = 4$, open triangles); b, worker honeybee *Apis mellifera* ($n = 3$, open circles), bumblebee *Bombus lapidarius* ($n = 3$, open squares); c, bee-fly *Bombylius major* ($n = 4$, filled circles), *Eristalis tenax* ($n = 3$, filled squares), hoverfly *Volucella pellucens* ($n = 3$, filled diamonds), blowfly *Calliphora erythrocephala* ($n = 3$, open triangles). Where range bars are shown, curves are the means ($\pm$ s.e.m.) from several recordings from cells with similar tuning. The stimulus intensity was as close as possible to that encountered in the normal habitat of each species (see below). Only hovering species of Diptera and Lepidoptera (filled symbols) respond strongly to temporal frequencies below 0.5 Hz, but hovering Diptera have wider response ranges than do the Lepidoptera.

METHODS. Cells were adapted to a blank screen (mean luminance) for at least 5 s before stimulation with a drifting, sinusoidal grating at one of 20 different temporal frequencies and at a spatial frequency close to optimum for each neuron (either 0.05 or 0.1 cycles deg^{-1}), determined from a preliminary experiment. Responses are normalized mean spike rates (spontaneous activity subtracted) taken from peristimulus time histograms in 10-ms bins. The effects of motion adaptation¹¹⁻¹³ on curve shape were minimized by using short stimulus duration and non-saturating stimulus contrast (Michelson contrast approximately 0.1). Under these conditions, peristimulus histograms indicated no response adaptation in most species. In hoverflies a small transient was visible, but normalized tuning curves derived from transient and sustained response levels were indistinguishable. Data for all species except the hawkmoth *Deilephila* (the only nocturnal species) were obtained during the day with unattenuated screen intensity (60 cd m^{-2} , equivalent to daylight and shade) and a 500-ms stimulus duration. *Deilephila* was studied at night using neutral-density filters (Kodak Wratten) to attenuate screen intensity to $6 \times 10^{-4} \text{ cd m}^{-2}$ (equivalent to partial moonlight), and using a longer stimulus duration of 2 s because the response time course is unusually slow in this species.

tered during hovering? For a periodic grating of spatial frequency f_s moving at velocity V ,

$$V = f_t / f_s \quad (1)$$

where f_t is the temporal frequency with which stripes pass a fixed point in space. To cope with low velocities, motion detectors must be sensitive to low temporal frequencies or to high spatial frequencies. The correlator delay mechanism (probably a low-pass filter, Fig. 1) produces a characteristic tuning to an optimum temporal frequency¹⁻³. We investigated whether hovering insects use motion detectors with longer delays, thus using lower temporal frequencies to code the lower velocities experienced during hovering.

Sphingid moths (hawkmoths) are the 'hummingbirds' of the insect world, hovering with great precision while they insert their long proboscis into flowers¹⁵. Sphingid neurons have the lowest temporal frequency optima of the insects studied: 0.7 Hz in the nocturnal species *Deilephila* (under fully dark-adapted conditions), and 2 Hz in a diurnal species, *Hemaris* (light adapted) (Fig. 2a). This low-frequency tuning suggests that their motion detectors use the longest correlator delays (Fig. 1).

In agreement with published data for other species^{13,16,17}, neurons in butterflies and bees respond optimally in the range 10–25 Hz (Fig. 2a, b). This higher-frequency tuning allows these fast-flying insects to respond vigorously to temporal frequencies beyond the range of human vision¹⁸⁻²¹. The bumblebee *Bombus*, which is unsteady at low speeds, has the highest optimum temporal frequency (25 Hz) of the ten species studied (Fig. 2b).

Responses of neurons from four species of dipteran flies peak at intermediate temporal frequencies (6–10 Hz). Bombyliid and syrphid flies hover for long periods while feeding from flowers or defending a territory, and their neurons are more responsive to low temporal frequencies (<0.5 Hz) than those of the blowfly *Calliphora* (Fig. 2c). Nevertheless, hovering flies appear to have higher temporal frequency optima than their moth counterparts.

To relate these temporal differences to image velocities, we need to account for the spatial tuning of motion correlators (equation (1)), as determined by the optics of the eye and the angular separation of the correlator inputs³. To see how these spatial properties interact with temporal frequency tuning to determine the velocity sensitivity of motion detectors (equation (1)), we plotted the contrast sensitivity at 200 combinations of spatial and temporal frequency (representing more than four log units of pattern velocity), in neurons from three species (Fig. 3). To our knowledge, such complete spatio-temporal plots have only previously been obtained from human observers¹⁸.

These spatio-temporal sensitivity measurements show that the large differences in temporal frequency tuning do indeed shift the range of velocities to which neurons are sensitive. Neurons in the hawkmoth *Deilephila* respond reliably to image velocities in the range 1–200 deg s⁻¹, with an optimum of 40 deg s⁻¹ (Fig. 3a), whereas bumblebee neurons respond to much higher velocities, from approximately 5 to above 2,000 deg s⁻¹, with an optimum of 350 deg s⁻¹ (Fig. 3b). Because neurons of both insects have similar spatial sensitivities (peaking at around 0.05 cycles deg⁻¹), and reliably detect a similar minimum image contrast of approximately 0.05, temporal tuning is the primary determinant of these differences in velocity sensitivity.

The hoverfly *Volucella* is very sensitive to contrast, detecting levels below 0.01 at an optimum spatial frequency that is nearly three times higher than in the bumblebee and hawkmoth (Fig. 3c). A higher sensitivity to contrast and better spatial resolution allows *Volucella* to respond to velocities above 2,000 deg s⁻¹, like the bumblebee, yet retain sensitivity to velocities below 0.5 deg s⁻¹, which is even slower than in the hawkmoth. This alternative strategy—extending sensitivity to low velocities by improving optical and neural resolution—accounts for the apparent discrepancy between hovering flight and temporal tuning in this species, and results in an optimum velocity similar to that in the hawkmoth (approximately 40 deg s⁻¹). A wide velocity range is

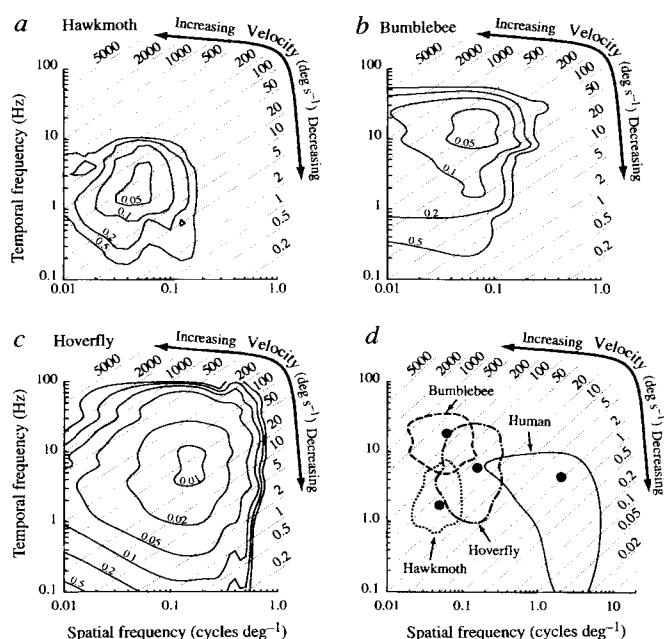


FIG. 3 a–c, Spatio-temporal contrast sensitivity of neurons from three species. Data are contour plots of threshold contrast (the stimulus contrast required to evoke a significance response). Two hovering species (a, hawkmoth and c, hoverfly) are both tuned to similar (low) image velocities. In the hawkmoth *Deilephila* (a) this results from a low temporal frequency optimum, but in the hoverfly *Volucella* (c) it results from a high spatial optimum. Despite similar spatial properties to *Deilephila*, the relatively clumsy bumblebee *Bombus* (b) has a much higher temporal optimum, and hence a higher velocity optimum. d, Optimum contrast sensitivity (dots) and the locus at which it falls to 50% of the normalized maximum (lines) for the three insects and for human vision (derived from published equations¹⁸). The temporal tuning of human motion detection is similar to that of the insects, but the optimum is at much higher spatial frequency and thus at lower velocities.

METHODS. We developed a rapid method to determine contrast sensitivity from 200 combinations of spatial and temporal frequency during single intracellular recordings from each neuron. Cells were adapted to a blank screen for 10 s. The contrast of a drifting grating was then increased linearly for 10 s, over the contrast range 0–1.0. The shape of the smoothed instantaneous spike response measured during a contrast ramp at near-optimal spatial and temporal frequency approximated the relation between spike rate and contrast obtained by conventional means, with drifting gratings presented at several different, constant contrast levels, but was shifted laterally by an amount corresponding to the response latency. After correcting for this shift (assumed constant for all stimulus combinations), the contrast required for the ramp response to exceed the mean background rate by 1 s.d. (usually close to 20% of the maximum response level) was taken as the threshold.

particularly appropriate for hoverflies such as *Volucella* which, after hovering for extended periods, may suddenly dart towards flowers or conspecifics at high velocity²². However, in *Volucella* this ability is achieved at the expense of compound eyes that are 4 mm in diameter, which makes them among the largest found in insects.

The relation observed here between ambient velocity and the spatio-temporal tuning of motion detectors is consistent with psychophysical data obtained for human motion detection¹⁸⁻²¹. Like flying insects, humans can detect high image velocities, but only for patterns with low spatial frequency, because the detection of moving patterns breaks down at temporal frequencies above 30–40 Hz (refs 18–21). However, the maximum sensitivity of the human spatio-temporal contrast sensitivity function¹⁸ lies at a high spatial frequency of approximately 2 cycles deg⁻¹ (Fig. 3d), and results in a low optimum velocity of approximately 2 deg s⁻¹. This is close to the velocity induced by natural drift movements of the eyes when fixating stationary objects¹⁸, the dominant motion experienced by the human visual system. A similarly 'slow'

optimum velocity would be desirable for visually mediated flight stabilization during hovering, when images are nearly stationary. However, optical constraints to the spatial acuity of insect compound eyes mean that a lower velocity optimum would not be attainable without impractically large correlation delays. Nevertheless, it seems that in insects, as in humans, the neural mechanisms for motion detection are matched, as far as possible, to motion experienced during behaviour. □

Received 27 February; accepted 21 May 1996.

1. Egelhaaf, M., Borst, A. & Reichardt, W. *J. opt. Soc. Am. A* **6**, 1070–1087 (1989).
2. Borst, A. & Egelhaaf, M. *Trends Neurosci.* **12**, 297–306 (1989).
3. Buchner, E. in *Photoreception and Vision in Invertebrates* (ed. Ali, M. A.) 561–621 (Plenum, New York, 1984).
4. Hassenstein, B. & Reichardt, W. *Z. Naturf.* **B 11**, 513–524 (1956).
5. Reichardt, W. *Z. Naturf.* **B 12**, 448–457 (1957).
6. Reichardt, W. in *Principles of Sensory Communication* (ed. Rosenblith, W. A.) 303–317 (Wiley, New York, 1961).
7. Barlow, H. B. & Levick, W. R. *J. Physiol., Lond.* **178**, 477–504 (1965).
8. Wolf-Oberhollenzer, F. & Kirschfeld, K. *J. Neurophysiol.* **71**, 1559–1573 (1994).
9. Wilson, H. R. *Biol. Cybern.* **51**, 213–222 (1985).
10. Santen, J. P. H. van & Sperling, G. *J. opt. Soc. Am. A* **2**, 300–321 (1985).
11. Maddess, T. & Laughlin, S. B. *Proc. R. Soc. Lond. B* **225**, 251–275 (1985).
12. Ruyter van Steveninck, R. de, Zaagman, W. H. & Mastebroek, H. A. K. *Biol. Cybern.* **54**, 223–236 (1986).
13. Maddess, T., DuBois, R. A. & Ibbotson, M. R. *J. exp. Biol.* **161**, 171–199 (1991).
14. Hausen, K. & Egelhaaf, M. in *Facets of Vision* (eds Stavenga, D. & Hardie, R. C.) 391–444 (Springer, Berlin, 1989).
15. Farina, W. M., Varju, D. & Zhou, Y. *J. comp. Physiol. A* **174**, 239–247 (1994).
16. Ibbotson, M. R. & Goodman, L. J. *J. exp. Biol.* **148**, 255–279 (1990).
17. Bidwell, N. J. & Goodman, L. J. *Apidologie* **24**, 333–354 (1993).
18. Kelly, D. H. *J. opt. Soc. Am. A* **69**, 1340–1349 (1979).
19. Burr, D. C. & Ross, J. *Vision Res.* **22**, 479–484 (1982).
20. Anderson, S. J. & Burr, D. C. *Vision Res.* **25**, 1147–1154 (1985).
21. Morgan, M. J. & Castet, E. *Nature* **378**, 380–383 (1995).
22. Collett, T. S. & Land, M. F. *J. comp. Physiol. A* **99**, 1–66 (1975).

ACKNOWLEDGEMENTS. We thank H. Barlow, D. Osorio and J. Anderson for comments and suggestions. R. Steel for donation of hive bees during winter. Supported by the EPSRC, BBSRC, The Swedish Natural Science Research Council and the British Council. N.J.B. was supported by a training fellowship from the Wellcome Trust.

CORRESPONDENCE and requests for materials should be addressed to D.C.O. (e-mail: dco1000@cam.ac.uk).

Abnormal processing of visual motion in dyslexia revealed by functional brain imaging

Guinevere F. Eden*, John W. VanMeter†, Judith M. Rumsey‡, José Ma. Maisog*, Roger P. Woods§ & Thomas A. Zeffiro||†

* Section on Functional Brain Imaging, NIMH, ‡ Child Psychiatry Branch, NIMH, || Laboratory of Diagnostic Radiology Research, OD, National Institutes of Health, Bethesda, Maryland 20892, USA
† Sensor Systems Inc., 103A Carpenter Drive, Sterling, Virginia 20164, USA

§ Department of Neurology, University of California at Los Angeles School of Medicine, Los Angeles, California 90098, USA

It is widely accepted that dyslexics have deficits in reading and phonological awareness^{1,2}, but there is increasing evidence that they also exhibit visual processing abnormalities that may be confined to particular portions of the visual system^{3,4}. In primate visual pathways, inputs from parvocellular or magnocellular layers of the lateral geniculate nucleus remain partly segregated in projections to extrastriate cortical areas specialized for processing colour and form versus motion^{5–10}. In studies of dyslexia, psychophysical¹ and anatomical⁴ evidence indicate an anomaly in the magnocellular visual subsystem. To investigate the pathophysiology of dyslexia, we used functional magnetic resonance imaging (fMRI) to study visual motion processing in normal and dyslexic men. In all dyslexics, presentation of moving stimuli

failed to produce the same task-related functional activation in area V5/MT (part of the magnocellular visual subsystem) observed in controls. In contrast, presentation of stationary patterns resulted in equivalent activations in V1/V2 and extrastriate cortex in both groups. Although previous studies have emphasized language deficits, our data reveal differences in the regional functional organization of the cortical visual system in dyslexia.

In the non-human primate brain, cells from the parvocellular and magnocellular subdivisions of the lateral geniculate nucleus (LGN) exhibit distinctive anatomical and physiological properties¹¹, and their projections remain partly segregated beyond the primary visual cortex. Neurons in the magnocellular layers (M-cells) have a phasic response and higher luminance contrast sensitivity, whereas neurons in the parvocellular layers (P-cells) have a more sustained response and sharper tuning, often opponent, to colour. In humans, homologous M-cell and P-cell systems have been identified on the basis of anatomical¹², psychophysical^{13,14} and clinical evidence¹⁵. Further, evidence from functional neuroimaging studies has demonstrated regional functional specialization for visual motion processing in an extrastriate visual area (V5/MT)^{16–19} that is thought to be dominated by input from the magnocellular stream. A defect in the M-cell pathway at the level of the LGN might be expected to result in impaired motion stimulus processing. Such a deficit should therefore be measurable psychophysically as an impaired ability to discriminate different stimulus velocities, or physiologically as impaired activation in MT/V5 during perception of moving visual stimuli.

Data suggesting that dyslexics have a selective deficit in the M-cell system include both psychophysical³ and electrophysiological⁴ differences in contrast sensitivity. Furthermore, post-mortem inspection of the LGN in dyslexic brains has indicated decreases in magnocellular neuron size, without differences in parvocellular layers⁴. Behavioural observations that some dyslexics have poor temporal judgement²⁰, visual instability²¹, and higher coherent motion-detection thresholds²² are consistent with a selective deficit in the M-cell system, but no direct neurophysiological demonstration of these deficits has been made.

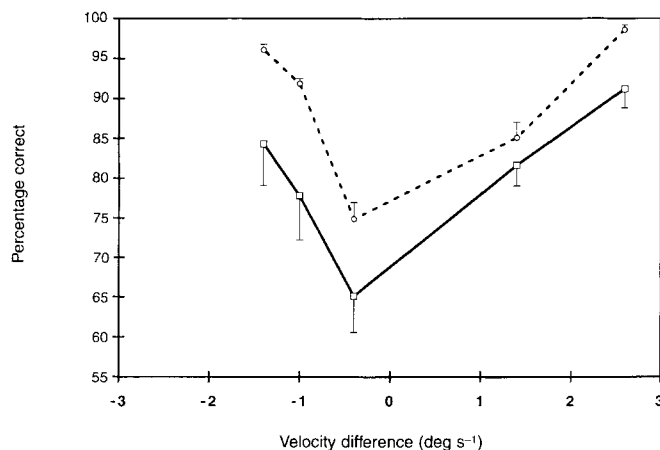


FIG. 1 Task performance is shown in a stimulus velocity detection task comparing the visual-motion sensitivity of six dyslexics (solid line) and six controls (broken line). The visual stimulus was the same as the M-stimulus used during fMRI data acquisition. Subjects saw low-contrast dots, moving horizontally with 100% coherence. Two such stimuli were presented in succession for 1 s each, the second differing from the first only in velocity. The subjects were instructed to state whether the second motion stimulus was slower or faster than the first. The first stimulus velocity was the same for all trials (7.0 deg s⁻¹) and the second stimulus velocity ranged from 5.6 to 9.6 deg s⁻¹. The dyslexics were significantly poorer in this task than the controls, showing reduced ability across the range of velocities tested (paired *t*-test, one-tailed, *P* < 0.03). The group mean for each stimulus pair is plotted with standard error bars.

We have tested the hypothesis that there is a selective M-cell-system processing deficit in dyslexia. Eight adult, male controls were carefully matched to six male dyslexics on IQ, age and other behavioural measures (Table 1). To assess whether there was a detectable visual-motion deficit in the dyslexic subjects, motion sensitivity was assessed by using a stimulus velocity judgement task. The results (Fig. 1) demonstrate a significant difference in performance between controls and dyslexics ($P < 0.03$).

Local blood-oxygenation level-dependent (BOLD^{23,24}) contrast signals were measured by using fMRI, while the subjects viewed either a coherently moving, low-contrast (5%), random-dot stimulus (M-stimulus), or a stationary, high-contrast (40%), patterned stimulus (P-stimulus). Signal changes were assessed as in a previous experiment¹⁹, by contrasting the motion condition to a fixation condition, and the pattern condition to a fixation condition. Because both the motion and the pattern stimulus were created from the same basic components (achromatic dark squares on a light background), we were able to equate the two comparisons in terms of their basic visual features. Area V5/MT was identified in each subject by its preferential sensitivity to motion (Fig. 2a).

Results of the single-subject analysis are shown in Table 2a. All control subjects exhibited bilateral motion sensitivity in a search volume surrounding the location of V5/MT¹⁷. In contrast, no activation was detected in any of the dyslexics in the same search volume, except for one subject's unilateral activation. A quite different result was observed with the stationary, pattern stimulus versus control comparison (Table 2b, c).

Subjects in both groups exhibited similar responses to the pattern stimulus in both posterior occipital cortex (V1/V2; Table 2b) and in extrastriate visual areas (inferior temporal/fusiform gyrus; Table 2c), demonstrating that the differences in visual processing between groups were both regional and stimulus specific.

A group analysis of variance (ANOVA) confirmed the results of the single-subject analysis. Figure 2b–e depicts *t* maps of the two visual stimulation conditions, each compared to the control stimulus for both groups. In normal subjects (Fig. 2b), motion sensitivity was found bilaterally in the V5/MT area, and additional motion-sensitive areas were seen outside the search volume, posterior and anterior to V5/MT. In the dyslexic group (Fig. 2c) there was no activation in the area V5/MT, and a reduced response in other motion-sensitive areas. However, a comparison of the stationary pattern versus control stimuli revealed similar activation patterns in V1/V2 for both control (Fig. 2d) and dyslexic (Fig. 2e) groups.

Our results confirm previous demonstrations of robust motion sensitivity in V5/MT in normal subjects^{17–19}. However, both individual and group analysis failed to detect responses to motion in dyslexics in this region, whereas normal responses to stationary, patterned stimuli were seen in V1/V2 and extrastriate visual areas. Further analysis confirmed that the activation differences between groups could not be attributed to higher signal intensities in V5/MT during the control condition. Because this study examined visual system function directly during a non-verbal condition, the observed differences could not be explained by the use of verbal strategies, which may be poorer in dyslexia².

Responses in other motion-sensitive, visual-cortical areas were detected in both groups, and raises the question of which mechan-

TABLE 1 Behavioural profile of subjects

	Controls (<i>n</i> = 8) (means (s.d.))		Dyslexics (<i>n</i> = 6) (mean (s.d.))		t-test
Chronological age (years)	25.5	(6.2)	26.8	(6.2)	n.s.
Education (highest level completed)	15.4	(2.7)	13.6	(1.4)	n.s.
Hollingshead's Socio-economic status	52.5	(27.6)	71.3	(21.9)	n.s.
Handedness	98.7	(3.3)	98.7	(3.3)	n.s.
Attention deficit disorder	0		0		n.s.
Wechsler adult intelligence scale-revised IQ					
Verbal	115.7	(16.4)	109.5	(9.1)	n.s.
Performance	106.7	(14.6)	115.5	(9.4)	n.s.
Full	112.7	(14.7)	114.0	(5.3)	n.s.
Wide range achievement test III					
single word reading	111.85	(4.9)	89.3	(11.8)	$P < 0.001$
spelling	107.4	(6.1)	69.8	(15.8)	$P < 0.001$
mathematics	112.7	(10.5)	93.8	(12.6)	$P < 0.005$
Gray oral reading test III					
rate	14.3	(1.4)	5.3	(2.9)	$P < 0.001$
accuracy	13.6	(2.4)	3.3	(2.2)	$P < 0.001$
passage score (decoding)	14.3	(1.8)	4.7	(2.4)	$P < 0.001$
Phonological skills (pseudo-word reading)	51.2	(4.8)	40.8	(2.6)	$P < 0.001$

All subjects were monolingual, adult American males, without neurological abnormalities. All were strongly right-handed (at least 92% on the handedness portion of the physical and neurological examination for subtle signs, and had no history of attention deficit disorder using the DSM-III-R criteria. Normal controls were closely matched to dyslexics on gender, age, education, socio-economic status and IQ. Reading was assessed with the Wide range achievement test (3rd edn) and the Gray oral reading test 3rd revision (GORT3). The dyslexic subjects met all of the following criteria: (1) a documented childhood history of reading disability; (2) an absolute reading deficit (below 8 points on the passage score of the GORT3); and (3) a discrepancy of at least 2 standard deviations between their reading (on the passage score of the GORT3) and verbal IQ (Wechsler adult intelligence scale-revised). The dyslexics demonstrated significantly poorer phonological awareness than controls on a pseudo-word (non-word) reading task (Goldman-Fristoe-Woodcock's reading of symbols). n.s., Not significant.

isms are responsible for the lack of activity in V5/MT. These temporal-lobe motion-sensitive areas have recently been reported to show particular sensitivity to coherent motion¹⁹, and their different onset times have been characterized using fMRI in conjunction with magnetoencephalography²⁵. The absence of motion sensitivity in V5/MT in dyslexics could result in a disruption of the normal coordinated temporal interaction with the other motion-processing areas. The results also raise the question of altered functioning of the non-geniculo-striate input to V5/MT (for example, from the superior colliculus and pulvinar) in dyslexia. Whatever the underlying mechanism, disruption of V5/MT activity may also interfere with re-entrant signals to other visual cortical areas (particular V1/V2), as well as to the oculomotor apparatus. This might explain the visual-motion-detection deficit demonstrated here, as well as previous reports of perceptual, oculomotor and evoked potential abnormalities in dyslexia^{4,21}.

In interpreting these findings, we must consider that, unlike patients with destructive lesions involving area V5/MT¹⁵ and surrounding structures, the motion-detection deficit of dyslexics is subtle. Careful psychophysical testing is required for its detection, and it is rarely severe enough to cause symptomatic complaint. The visual-motion deficit we observed in dyslexia resembles that seen in non-human primates after recovery from focal MT lesions. Although injection of ibotenic acid into V5/MT causes severe, acute motion-detection deficits, motion sensitivity recovers over a period of days to weeks. The animals eventually show only minimal impairment of motion detection, despite a complete absence of neurons in area V5/MT²⁶. Although focal ischaemic lesions of V5/MT can cause severe and permanent motion-detection deficits, it is possible that the ibotenic-acid

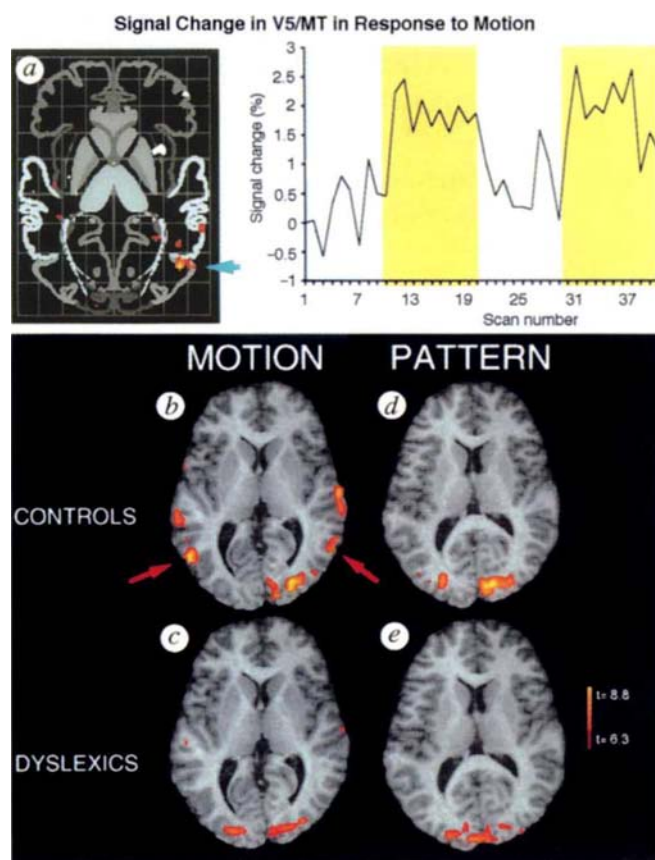
TABLE 2 Local voxel maxima

Subjects	Left				Right			
	Talairach coordinates (mm)				Talairach coordinates (mm)			
	x	y	z	z-score	x	y	z	z-score
(a) In V5/MT for motion versus fixation contrast								
Controls								
C1	-62	-72	+10	4.5	+58	-62	+12	5.6
C2	-64	-78	+10	5.6	+56	-68	+2	4.6
C3	-48	-84	+4	6.4	+34	-84	+2	7.6
C4	-50	-64	+8	5.3	+60	-62	+8	5.1
C5	-32	-74	+10	4.7	+52	-64	+6	4.2
C6	-60	-66	+10	5.3	+44	-78	+2	5.7
C7	-58	-74	-4	4.4	+52	-76	+2	5.7
C8	-44	-86	+12	6.9	+44	-66	+4	6.4
Dyslexics								
D1				n.s.				n.s.
D2				n.s.				n.s.
D3				n.s.				n.s.
D4				n.s.				n.s.
D5	-42	-62	+8	5.6				n.s.
D6				n.s.				n.s.
(b) In V1/V2 for pattern versus fixation contrast								
Controls								
C1	-12	-82	-8	7.6	+14	-90	-10	5.6
C2	-6	-82	0	4.8	+6	-96	+2	4.7
C3	-4	-76	+2	7.3	+22	-88	+14	6.1
C4	-24	-74	+6	6.0	+28	-74	+8	4.2
C5	-22	-92	+12	4.1	+6	-96	+10	4.9
C6	-22	-104	+4	4.1				n.s.
C7				n.s.	+18	-96	+4	4.3
C8	-14	-94	+8	5.1			ns	
Dyslexics								
D1	-16	-98	-2	5.9	+16	-94	-4	5.4
D2	-18	-88	+2	4.4				n.s.
D3	0	-105	+6	4.1	0	-105	+6	4.1
D4	-24	-70	-6	4.2	+6	-82	-6	5.5
D5	-2	-96	0	4.6	0	-106	+2	4.1
D6	-24	-80	+8	8.1	+6	-78	+10	6.5
(c) In extrastriate cortex (inferior temporal/fusiform gyrus) for pattern versus fixation contrast								
Controls								
C1	-58	-48	-16	6.1	+58	-48	-16	5.7
C2	-62	-42	-6	4.4	+46	-48	-6	5.1
C3	-48	-32	-10	6.4	+50	-40	-27	5.3
C4	-64	-48	-8	4.9	+64	-32	-8	4.0
C5				n.s.	+54	-48	-24	4.7
C6				n.s.				n.s.
C7	-46	-32	-14	4.0	+54	-32	-12	4.7
C8				n.s.				n.s.
Dyslexics								
D1	-64	-32	-14	5.4	+64	-32	-14	5.4
D2	-58	-49	-18	4.1				n.s.
D3	-40	-42	-6	4.1	+50	-32	-14	4.4
D4				n.s.				n.s.
D5	-52	-44	-16	4.3	+52	-42	-12	4.6
D6	-38	-44	-8	4.2	+48	-36	-24	4.5

Single-subject analysis, showing the locations in the Talairach atlas²⁷ for the most significantly activated voxels in response to: (a) the motion stimulus in V5/MT; (b) the stationary pattern stimulus in V1/V2; and (c) the stationary pattern stimulus in extrastriate object/form area²⁸ in normals and dyslexics. The time series for each of these voxels was examined to confirm that the signal changes were related to the task. As our interest was specific to area V5/MT, the search for areas sensitive to motion was confined to a volume containing the Talairach²⁷ coordinates reported to contain V5/MT¹⁷ (x, lateral to +30 or -30; y, posterior to -62; z, between -4 and +12 in the inferior/superior direction). The borders for V1/V2 were determined using the Talairach atlas²⁷. Spatial normalization of the functional images allowed accurate neuroanatomical localization for single-subject analysis as well as intersubject averaging for the group study. MEDx (Sensor Systems, Herndon, VA) was used for image analysis and visualization. The matrices from the following transformations were concatenated and applied to the original echo planar imaging (EPI) data²⁹. (1) The EPI data from each scan were co-registered with a 6 degrees-of-freedom (d.o.f.) transformation³⁰ to correct for interscan head motion. (2) Geometric spatial distortions caused by magnetic field inhomogeneities were corrected with a 12 d.o.f. transformation by registering the EPI images to a coplanar conventional volume, acquired during the same session (spoiled grass gradient recalled echo (SPGR): field of view (FOV), 32 cm; acquisition matrix, 256 × 256; slice thickness, 5.0 mm; echo time (TE), 9 ms; repetition time (TR), 150 ms; 30 contiguous coronal slices). (3) A high-resolution conventional scan was acquired on a separate day (SPGR: FOV, 24 cm; acquisition matrix, 512 × 512; slice thickness, 1.5 mm; TE, 5 ms; TR, 24 ms; 124 sagittal slices). High-resolution scans were spatially normalized into the Talairach space²⁷ by identifying the locations of the anterior and posterior commissures (AC-PC line) and applying a second-order polynomial transformation. (4) The coplanar SPGR volume was registered to the high-resolution SPGR volume. In this way EPI images were transformed into Talairach space, resulting in stereotactically normalized EPI data from which *t*- and *z*-score maps were computed.

FIG. 2 a, Left, a single control subject's brain after spatial normalization²⁷, and calculation of z-maps. Right, graph shows changes in image intensity in the most significantly activated, single voxel located in right area V5/MT. Activation for V5/MT is only displayed for the right hemisphere, as V5/MT on the left side lies in a more superior axial plane. The transverse image shows V5/MT located at Talairach coordinates $x = +44$, $y = -66$, and $z = +4$, which also corresponds to the most common anatomical location of V5/MT, at the ascending limb of the inferior temporal sulcus^{17,18}. b–e, For group comparisons, hypothesis testing was performed using an ANOVA and planned contrasts for each group in the two conditions. Statistical maps are displayed on one subject's spatially normalized MRI, shown at transverse planes $z = +8$ and $z = +10$ relative to the AC–PC line, for the motion stimulus versus control (b, c) and pattern stimulus versus control (d, e) contrasts. V5/MT can clearly be seen in the control group (b) but not in the dyslexic group (c) by its preferential activation in the moving-stimulus condition. In contrast, both groups (d, e) show similar responses in area V1/V2 to the stationary, high-contrast, patterned stimulus. In the motion condition, additional motion-sensitive areas can be seen in the control group (b) at anterior temporal as well as posterior areas. Neuroimaging studies of motion sensitivity have proposed the latter area to be V3 (refs 18, 19). Both of these areas were also identified, to a lesser extent, in the dyslexic group. All images are presented in the neurological convention (subject left, image left).

METHODS. Data are derived from multislice EPI data of local BOLD^{23,24} contrast signals, on a GE Signa 1.5 Tesla system with two surface coils positioned in an oblique orientation surrounding the occiput. Contiguous coronal 5-mm slices (30), 32 cm; acquisition matrix, 64×64 ; TE, 40 ms; TR, 10 s) were collected, resulting in a volume of 5-mm cubic voxels that spanned occipital and posterior parietal cortex; 90 scans were obtained for each subject. During the scans, subjects maintained fixation while viewing: (1) a fixation cross on a field of uniform illumination; (2) a low-contrast (Michaels contrast, 5%), array of black dots on a grey background, all moving at 100% coherence in one of eight directions at a speed of 10 deg s^{-1} (M-stimulus); or (3) a high-contrast (40%), stationary, patterned dot field with high spatial correlation, resulting in texture (P-stimulus). All three stimuli had equal luminance, moving and stationary pattern stimuli had the same global spatial frequency. Each dot (square) subtended $1'$. All subjects reported perceiving motion in (2). Stimuli were presented on a back-projection screen, and the display covered 30° of the horizontal and 20° of the vertical central visual field. Blocks of 10 trials of viewing one of the three stimuli were presented, and repeated to yield 90 scans (123123123). A 1-min interscan interval minimized temporal autocorrelation and neural habituation. After ratio normalization, task-related signal changes were detected by computing voxel-wise t-statistic maps contrasting motion versus fixation to detect motion sensitivity, and stationary



pattern versus fixation to detect pattern sensitivity. After transformation to the standard normal distribution, areas exhibiting motion or pattern sensitivity were identified using z-score ≥ 4 ($P < 0.00003$) as the critical threshold. After correction for multiple comparisons with reference to the search volume used, the critical threshold for the V5/MT volume was $P < 0.003$, and the threshold for the V1/V2 volume was $P < 0.01$.

lesion is a better model of the developmental lesion that may result in the visual-motion detection impairment observed in dyslexia.

We have demonstrated the feasibility of using fMRI to detect and localize abnormal neuronal processing in dyslexia. Our findings provide a neurophysiological basis for the previously observed visual perceptual processing deficits in dyslexia that have implicated the M-cell system. These visual-system abnorm-

alities may be one component, and therefore a marker, of a disorder that embodies numerous constituents, including the well-studied deficit in phonological awareness^{1,2}. Their joint appearance in dyslexia may be due to the presence of an underlying deficit in systems that have in common the processing of temporal properties of stimuli. This deficit may manifest itself as disorders of phonological awareness, rapid naming, rapid visual processing, or motion detection. □

Received 26 January; accepted 17 May 1996.

- Bradley, L. & Bryant, P. *Nature* **271**, 746–747 (1978).
- Snowling, M. *Psychol. Res.* **43**, 219–234 (1981).
- Lovegrove, W. J., Bowling, A., Badcock, D. & Blackwood, M. *Science* **210**, 439–440 (1980).
- Livingstone, M., Rosen, G. D., Drislane, F. W. & Galaburda, A. *Proc. natn. Acad. Sci. U.S.A.* **88**, 7943–7947 (1991).
- Zeki, S. M. *Brain Res.* **53**, 422–427 (1973).
- Ungerleider, L. G. & Mishkin, M. in *Analysis of Visual Behavior* (eds Ingle, D. J., Goodale, M. A. & Mansfield, R. J. W.) 549–586 (MIT Press, Cambridge, MA, 1982).
- Livingstone, M. & Hubel, D. *Science* **240**, 740–749 (1988).
- Maunsell, J. H. R. & Newsome, W. A. *Rev. Neurosci.* **10**, 363–401 (1987).
- Van Essen, D. & Maunsell, J. H. R. *Trends Neurosci.* **6**, 370–375 (1983).
- Zeki, S. M. & Shipp, S. *Nature* **335**, 311–317 (1988).
- Shapley, R. & Perry, V. *Trends Neurosci.* **9**, 229–235 (1986).
- Clarke, S. & Miklosy, J. *J. comp. Neurol.* **298**, 188–214 (1990).
- Kaplan, E. & Shapley, R. *J. Physiol., Lond.* **330**, 125–143 (1982).
- Kulikowski, J. & Tolhurst, D. J. *J. Physiol., Lond.* **232**, 149–162 (1973).
- Zihl, J. *Brain* **106**, 313–340 (1983).
- Corbetta, M., Miezin, F. M., Dombeyer, S., Shulman, G. L. & Petersen, S. E. *Science* **248**, 1556–1559 (1990).
- Watson, J. D. G. et al. *Cerebr. Cortex* **3**, 79–94 (1993).
- Tootell, R. et al. *J. Neurosci.* **15**, 3215–3230 (1995).

- Cheng, K., Fujita, H., Kanno, I., Miura, S. & Tanaka, K. *J. Neurophysiol.* **74**, 413–427 (1995).
- Eden, G., Stein, J., Wood, H. & Wood, F. *Cortex* **31**, 451–469 (1995).
- Eden, G., Stein, J., Wood, H. & Wood, F. *Vision Res.* **34**, 1345–1358 (1994).
- Cornelissen, P., Richardson, A., Mason, A., Fowler, S. & Stein, J. *Vision Res.* **35**, 1483–1494 (1995).
- Ogawa, S. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 5951–5955 (1992).
- Kwong, K. K. et al. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5676–5679 (1992).
- Dale, A. et al. *Soc. Neurosci. Abstr.* **21**, 10 (1995).
- Newsome, W. T. & Paré, F. B. *J. Neurosci.* **8**, 2201–2211 (1988).
- Talairach, J. & Tournoux, P. *Co-planar Stereotaxic Atlas of the Human Brain* (Thieme, New York, 1988).
- Sergent, J., Ohta, S. & MacDonald, B. *Brain* **115**, 15–36 (1995).
- Zeffiro, T. A., Eden, G. F., Woods, R. P. & vanMeter, J. W. *Adv. exp. med. Biol.* (in the press).
- Woods, R. P., Cherry, S. R. & Mazziotta, J. C. *J. Comput. assist. Tomogr.* **16**, 620–633 (1992).

ACKNOWLEDGEMENTS. We thank P. Jezzard, J. Haxby and L. Ungerleider for discussions, and C. Moonen and the NIH in vivo NMR Center for use of their facilities. This work was supported by a grant from the Human Science Frontier Program. G.F.E. was supported by a Fogarty fellowship, sponsored by Judith Rapoport.

CORRESPONDENCE and requests for materials should be addressed to G.F.E. (e-mail: eden@alw.nih.gov).

Renal agenesis and the absence of enteric neurons in mice lacking GDNF

Marina P. Sánchez*, Inmaculada Silos-Santiago, Jonas Frisén, Bin He, Sergio A. Lira† & Mariano Barbacid

Department of Molecular Oncology and † Transgenic Unit, Bristol-Myers Squibb Pharmaceutical Research Institute, Princeton, New Jersey 08543, USA

* Present address: Centro de Biología Molecular, Universidad Autónoma, Madrid, 28049 Spain.

GLIAL-CELL-LINE-DERIVED neurotrophic factor (GDNF)¹ is a potent survival factor for dopaminergic neurons and motor neurons in culture^{1,2}. It also protects these neurons from degeneration *in vitro*³⁻⁹, and improves symptoms like Parkinson's disease induced pharmacologically in rodents¹⁰ and monkeys¹¹. Thus GDNF might have beneficial effects in the treatment of Parkinson's disease and amyotrophic lateral sclerosis. To examine the physiological role of GDNF in the development of the mammalian nervous system, we have generated mice defective in *GDNF* expression by using homologous recombination in embryonic stem cells to delete each of its two coding exons¹. *GDNF*-null mice, regardless of their targeted mutation, display complete renal agenesis owing to lack of induction of the ureteric bud, an early step in kidney development. These mice also have no enteric neurons, which probably explains the observed pyloric stenosis

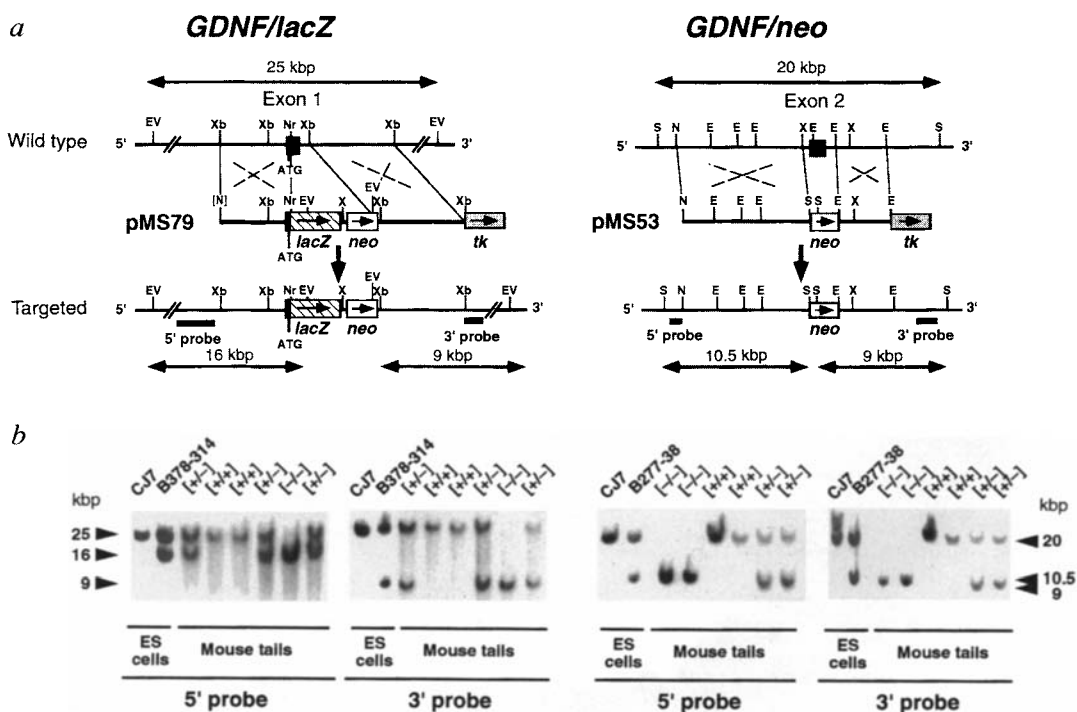
and dilation of their duodenum. However, ablation of the *GDNF* gene does not affect the differentiation and survival of dopaminergic neurons, at least during embryonic development.

Two strains of mice carrying germline mutations in the *GDNF* locus were generated by blastocyst injection of recombinant embryonic stem (ES) cell clones containing either a bacterial *lacZ* gene fused in frame with the first eight coding nucleotides of the *GDNF* gene (*GDNF/lacZ* mice), or a phosphoglycerate kinase-1/neo cassette (*PGK-neo*), replacing the entire second coding exon of *GDNF* (*GDNF/neo* mice) (Fig. 1). Heterozygous *GDNF/lacZ* mice were used for analysis of the pattern of expression of *GDNF* during embryonic development. In agreement with previous studies, we found *GDNF* to be expressed in embryonic regions where inductive mesenchymal-epithelial interactions are known to occur¹². Detectable X-gal staining appears for the first time in embryos at embryonic day 8.5 (E8.5) in the nephrogenic cord, primitive gut and the first branchial pouch. At E10.5, *lacZ* expression can be observed in the developing nervous system, specifically in the ventral regions of the midbrain and hindbrain, and in ventral and dorsal regions of the mid/hindbrain junction. In addition, X-gal staining could be observed in the craniofacial mass, the umbilical gut, the mesonephros, and the most proximal regions of both fore- and hindlimbs. At E14.5, additional X-gal staining was observed in the developing mystacial pad, the intercostal and shoulder muscles, and the metanephros (data not shown).

Crosses of heterozygous *GDNF/lacZ* or *GDNF/neo* mice resulted in approximately the expected mendelian ratio of wild-type (53 of 177, 30%), heterozygous (90 of 177, 51%) and homozygous (34 of 177, 19%) animals, indicating that *GDNF* expression is not absolutely required for embryonic development. However, all *GDNF*-null mice, regardless of the nature of the

FIG. 1 Generation of *GDNF* mutant mice. **a**, Schematic representation of the targeting strategy used to mutate the first (left, *GDNF/lacZ* mice) or second (right, *GDNF/neo* mice) coding exons of the mouse *GDNF* locus. Black boxes represent *GDNF* coding exons. The bacterial *lacZ* gene fused in frame to the first eight coding nucleotides of *GDNF* is indicated by a hatched box, the PGK-1/neo cassette (*neo*) is indicated by a white box, and the PGK-1/HSV tk cassette (*tk*) by a grey box. Arrows indicate the direction of transcription. Vertical lines outline the *GDNF* genomic sequences incorporated into the targeting plasmids, pMS79 (left) and pMS53 (right). *EcoRI* (E), *EcoRV* (EV), *NotI* (N), *NruI* (Nr), *SpeI* (S), *XbaI* (Xb), *XhoI* (X). The diagnostic *EcoRV* (*GDNF/lacZ* mice) and *SpeI* (*GDNF/neo* mice) DNA fragments identified in wild-type (top) and targeted (bottom) DNAs are indicated. **b**, Southern blot analysis of DNA isolated from wild-type CJ7 ES cells²¹, ES cell clones targeted at the first (B378-314 cells) or second (B227-38 cells) *GDNF* coding exons and tails obtained from one-week-old mice derived from crosses of heterozygous *GDNF/lacZ* (left) or *GDNF/neo* (right) mice. The genotype of each mouse is indicated. Migration of the diagnostic DNA fragments is indicated by arrowheads.

METHODS. The targeting plasmids, pMS79 and pMS53, were generated by



subcloning the indicated mouse 129 *GDNF* genomic sequences into *DKS/lacZpA*, a targeting vector that contains the bacterial *lacZ* gene and pTNT²², respectively. Targeting plasmids were linearized and transfected into CJ7 ES cells as described²³. ES (*neo^R/gan^R*) clones displaying the expected recombinational events were microinjected into mouse blastocysts²⁴, and the resulting chimaeras mated to C57BL/6 mice until they transmitted the targeted allele.

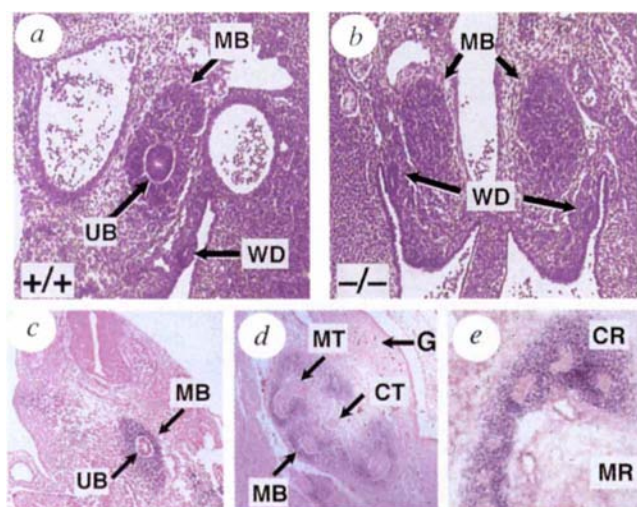


FIG. 2 The ureteric bud does not invade the metanephric blastema in *GDNF* null mice. *a*, In wild-type E11.5 embryos, the ureteric bud (UB) invades the metanephric blastema (MB). *b*, However, no ureteric bud can be observed in the blastemal region of *GDNF* null E11.5 embryos. Histological identification of *lacZ*-expressing cells²⁵ in transverse (*c*) or parasagittal (*d*, *e*) paraffin sections (8 μ m) through the urogenital region. *c*, In E10.5 embryos, intense X-gal staining is observed within the metanephric blastema (MB), but not in the developing ureteric bud (UB). *d*, In E12.5 embryos, X-gal staining is prominent in the condensed mesenchymal cells of the metanephros, but barely detectable in the metanephric tubules (MT), and is absent in the collecting tubes (CT). *e*, In E15.5 embryos, X-gal staining is localized throughout the cortical region (CR) but not in the medullar region (MR). WD, wolffian duct; G, gonad.

targeted mutation, die within their first postnatal day of life. Gross anatomical analyses of these mice indicate renal agenesis, pyloric stenosis, and dilated proximal small intestine, as well as an empty bladder. In contrast, the rest of the urogenital system, such as the gonads and adrenal glands, develops normally. *GDNF* is also highly expressed in the limb buds, structures derived from the branchial arches and pouches (including mandible, craniofacial mass and pharynx), and the brain; these regions seem to develop normally.

The development of the metanephric kidney is initiated around E11, when the ureteric bud branches from the wolffian duct into the metanephric blastema, a process thought to be triggered by an inductive signal sent by the mesenchymal-blastemal cells¹³. In *GDNF/lacZ*-null mice, the wolffian duct evolves normally along the length of the urogenital ridge in E9.5 and E10.5 embryos. Moreover, E11.5 homozygous embryos have a normal mesonephros and metanephric blastema. However, these mutant mice do not develop the ureteric bud, or if they do it never invades the metanephric mesenchyme (Fig. 2*a, b*). Robust levels of *GDNF* expression (as determined by X-gal staining) can be observed in the blastemal cells, but not in the ureteric bud from E10.5 *GDNF/lacZ*^{+/-} embryos (Fig. 2*c*). These results suggest that *GDNF* mediates the critical inductive signal required for budding and subsequent development of the ureteric bud.

In the absence of *GDNF*, the mesenchymal-blastemal cells undergo apoptosis, almost disappearing by E13.5. In heterozygous mice, robust *GDNF* expression remains in the condensed

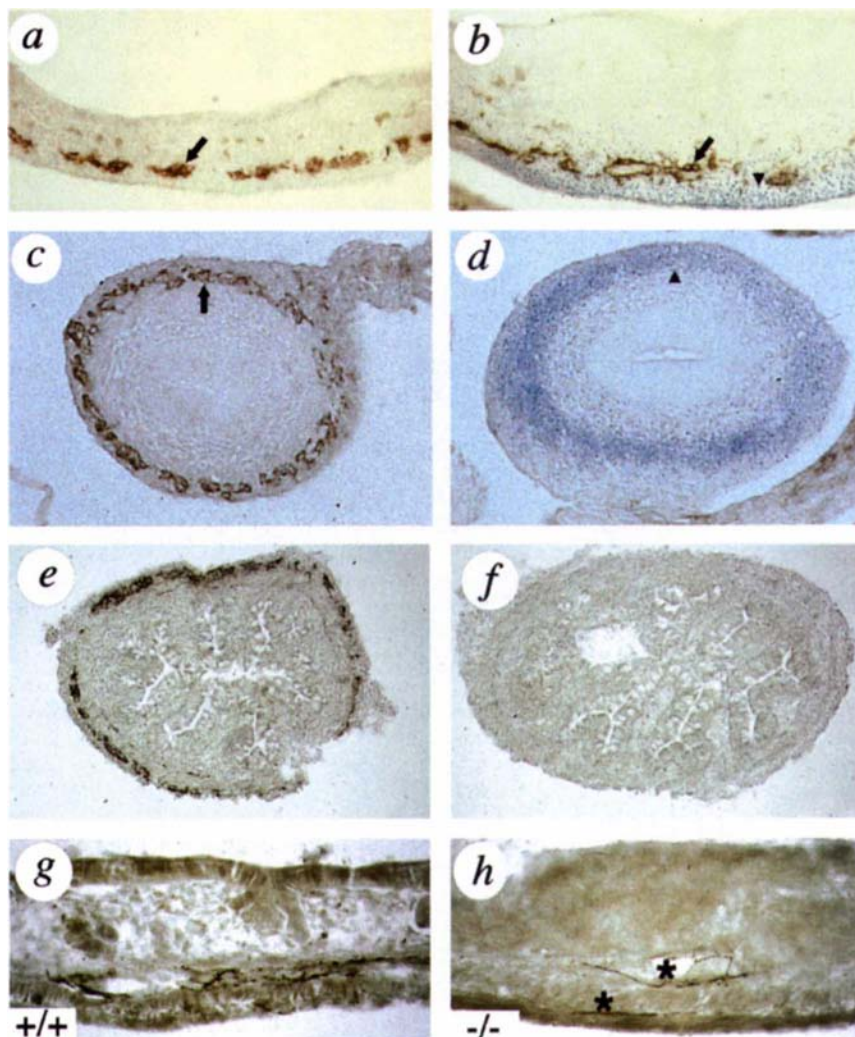


FIG. 3 ENS neurons do not survive in *GDNF* null mice. Transverse sections of wild-type (left column) and *GDNF* null (right column) E12.5 embryos (*a-d*) and newborn mice (*e-h*). Paraffin sections (8 μ m) of the stomach wall (*a*, *b*) and midgut intestinal loops (*c*, *d*) of the developing gut were stained with anti-p75 antibodies¹⁴. Frozen sections (14 μ m) through the gastrointestinal tract of newborn animals were stained either with AChE histochemistry²⁶ (*e*, *f*) or with anti-TH antibodies (Eugene Tech International (*g*, *h*)). Arrows denote p75-immunoreactive neurons. Arrowheads indicate regions of X-gal staining within the mesenchymal layer of the gastrointestinal wall. Asterisks indicate tyrosine hydroxylase immunoreactive fibres in *GDNF* null mice. Sections were examined on a Zeiss Axiophot.

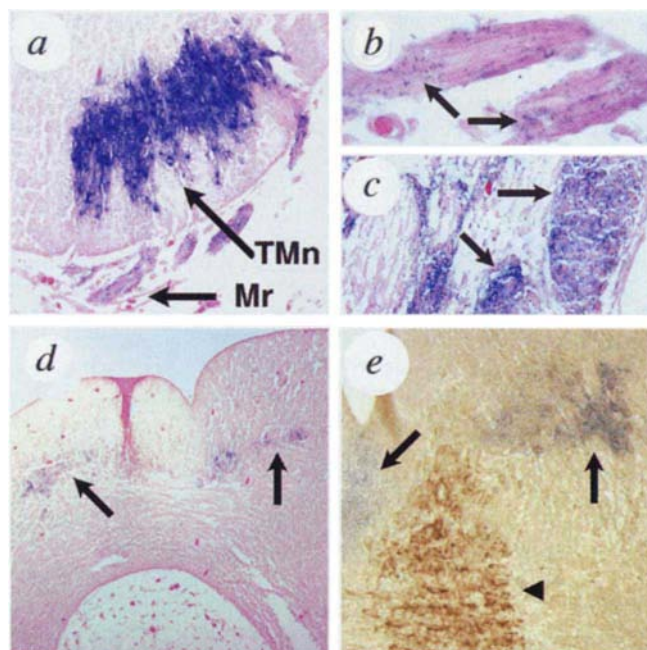


FIG. 4 *GDNF* expression in the hindbrain and midbrain of E12.5 embryos. Sections (8 μ m) depict the trigeminal motor nucleus (TMn) and motor roots (Mr) (a), ventral roots (b), skeletal muscle (c) and the A9–A10 region and floorplate of the midbrain (d, e). Cells displaying X-gal staining (arrows) (a–e) and tyrosine hydroxylase immunoreactivity (arrowhead) (e) are indicated.

mesenchyme until birth (Fig. 2d, e), suggesting that *GDNF* may be required throughout development for epithelial transformation of the condensed mesenchyme. However, very little *GDNF* expression could be found in the tubular epithelium, suggesting that *GDNF* expression is downregulated after epithelial differentiation. As expected, no X-gal staining could be observed in the collecting tubes that originate from the ureteric bud (Fig. 2d, e). *GDNF* heterozygous mice, regardless of the targeted mutation, occasionally exhibit either one (18%) or no (14%) kidneys. Analysis of these mice during embryonic development also indicated either unilateral or a complete lack of ureteric bud development. These observations indicate that the levels of *GDNF* produced by the blastemal cells are critical for the induction of metanephric development.

GDNF null mice display dilation of the proximal intestine and severe occlusion of the pyloric sphincter, suggesting a defect in those innervations that coordinate control of the gastrointestinal motility. To determine whether the enteric nervous system (ENS) is affected in these mice, we analysed embryos at those stages (E9.5 to E13.5) in which ENS precursors populate the gastrointestinal wall. For instance, at E12.5, antibodies against the p75 neurotrophin receptor¹⁴ revealed neurons in the oesophagus (not shown) and in the stomach wall (Fig. 3a, b), but not in the more distal midgut intestinal loops (Fig. 3c, d). In newborn animals, no ENS neurons could be detected in sections from the lower oesophagus, stomach, small intestine or colon, using either antibodies against nitric oxide synthase (not shown) or acetylcholinesterase (AChE) histochemistry (Fig. 3e, f). Another antibody, PGP 9.5 (Ultraclone), a general marker of neurons and their projections, also failed to detect any ENS neurons in the *GDNF* null mice (results not shown). However, this antibody labelled a few fibres in the gastrointestinal wall of these mutant mice. Most, if not all, of these fibres were positive for tyrosine hydroxylase staining, suggesting that they correspond to sympathetic innervation (Fig. 3g, h). These results suggest that, although some ENS neurons can develop in the absence of *GDNF*, they die during late embryonic development.

GDNF is expressed in the area where dopaminergic neurons develop (A9–A10 regions) and in the floorplate (Fig. 4) and to a limited extent in striatum, a major target for midbrain dopaminergic neurons. However, no defects were observed in the number of tyrosine hydroxylase-positive, dopaminergic neurons in these regions ($13,771 \pm 499$ in (+/+) versus $15,010 \pm 424$ in (-/-) mice, $n = 3$). Similarly, tyrosine hydroxylase immunohistochemistry in sections throughout the basal ganglia failed to detect any alterations in the innervation pattern of these dopaminergic neurons in the *GDNF* mutant mice. Furthermore, no differences were observed in the number of neurons within the locus coeruleus (796 ± 34 in (+/+) versus 770 ± 37 in (-/-) mice, $n = 5$), a noradrenergic nucleus that is severely impaired in neurodegenerative disorders, and whose neurons are rescued by *GDNF* from chemical-induced degeneration¹⁵.

The trigeminal motor nucleus, a region where *GDNF* is highly expressed (Fig. 4a), displays a 21% decrease in motor neuron number in *GDNF* null mice, regardless of the nature of the mutation (899 ± 44 in (+/+) versus 711 ± 11 in (-/-) mice; $n = 6$). The L4–L5 levels of the spinal cord of these mutant mice also have a 31% reduction in their motor neurons ($2,312 \pm 88$ in (+/+) versus $1,592 \pm 47$ in (-/-) mice; $n = 5$). Although *GDNF* is not expressed in these neurons or in the surrounding glia (not shown), high levels of X-gal staining can be observed in the ventral motor roots and in skeletal muscle (Fig. 4b, c). No significant differences were found in the number of facial motor neurons ($3,576 \pm 187$ in (+/+) versus $3,422 \pm 95$ in (-/-) mice; $n = 5$), despite the protective effect that *GDNF* exerts on these neurons after axotomy^{3,7,9}. Further studies will be needed to determine the requirements of *GDNF* for the survival of different populations of motor neurons.

Significant levels of X-gal staining were also observed in other regions of mesenchymal–epithelial interactions in *GDNF*^{+/+} mice, including the nasal cavity and developing face cartilage, developing teeth, primordial whisker follicles, limb bud, and interdigital mesenchyme (not shown). Other regions depicting *lacZ* expression include skin, proximal muscles of the developing limbs, as well as intercostal mesenchyme and muscles. No defects were observed in any of these regions in the *GDNF* null mice.

Mice defective for the gene encoding the Ret tyrosine protein kinase receptor show remarkable similarities with those lacking *GDNF*¹⁶. These *ret* homozygous mutants do not form the ureteric bud or undergo metanephric development, and lack ENS neurons¹⁶. These observations, along with the complementary expression patterns of *GDNF* and Ret proteins¹⁷, strongly suggest that Ret is a signalling receptor for *GDNF*. Although superior cervical ganglion (SCG) neurons do not survive in *ret*^{-/-} mice¹⁸, the SCG of *GDNF* null mice only display a limited reduction (20%) in volume. These observations suggest that the Ret receptor can be activated by other growth factors.

Other strains of knockout mice have also shown defects in kidney development similar to those observed in the *GDNF* null mice. Newborn mice homozygous for a targeted mutation in the transcriptional regulator *Pax-2* lack kidneys, ureters and genital tracts¹⁹, probably as a result of dysgenesis of the mesenchymal and ductal components of the urogenital system, structures in which *Pax-2* is expressed¹⁹. In mice defective for WT-1, a tumour-suppressor gene that encodes another transcription factor, cells of the metanephric blastema undergo apoptosis at about E11, thus preventing induction of the ureteric bud and subsequent development of the metanephric kidney²⁰. These results suggest that *GDNF* expression might be regulated by these transcription factors.

Here we have demonstrated the pleiotropic roles of *GDNF* during embryonic development. Ectopic expression of *GDNF* in kidney and gastrointestinal tissue of transgenic mice may overcome the neonatal lethality of the *GDNF* null mice. Such animals should help to determine whether *GDNF* also serves as a neurotrophic factor during postnatal development and in adult mice. □

Received 11 April; accepted 2 June 1996.

- Lin, L. F. H. et al. *Science* **260**, 1130–1132 (1993).
- Henderson, C. E. et al. *Science* **226**, 1062–1064 (1994).
- Zum, A. D. et al. *NeuroReport* **6** (1), 113–118 (1994).
- Beck, K. D. et al. *Nature* **373**, 339–341 (1995).
- Kearns, C. M. & Gash, D. M. *Brain Res.* **672**, 104–111 (1995).
- Li, L. X. et al. *Proc. natn. Acad. Sci. U.S.A.* **92**, 9771–9775 (1995).
- Oppenheim, R. W. et al. *Nature* **373**, 344–346 (1995).
- Tomac, A. et al. *Nature* **373**, 335–339 (1995).
- Yan, Q., Matheson, C. & Lopez, O. T. *Nature* **373**, 341–344 (1995).
- Hoffer, B. J. et al. *Neurosci. Lett.* **182** (1), 107–111 (1994).
- Gash, M. D. et al. *Nature* **380**, 252–255 (1996).
- Hellmich, H. L. et al. *Mech. Dev.* **54**, 95–105 (1996).
- Saxen, L. *Organogenesis of the Kidney* Vol. 19 (eds Barlow, P. W., Green, P. B., and Wylie, C. C.) (Cambridge Univ. Press, 1987).
- Blaugrund, E. et al. *Development* **122**, 309–320 (1996).
- Arenas, E., Trupp, M., Akerud, P. & Ibañez, C. *Neuron* **15**, 1465–1473 (1995).
- Schuchardt, A. et al. *Nature* **367**, 380–383 (1994).
- Pachnis, V., Mankoo, B. & Costantini, F. *Development* **119**, 1005–1017 (1993).
- Durbec, P. L. et al. *Development* **122**, 349–358 (1996).
- Torres, M., Gomez-Pardo, E., Dressler, G. R. & Gruss, P. *Development* **121**, 4057–4065 (1995).
- Kreidberg, J. A. et al. *Cell* **74**, 679–691 (1993).
- Swiatek, P. J. & Gridley, T. *Genes Dev.* **7**, 2071–2084 (1993).
- Tybulewicz, V. L. et al. *Cell* **65**, 1153–1163 (1991).
- Wurst, W. & Joyner, A. L. in *Gene Targeting: a Practical Approach* (ed. Joyner, A. L.) 33–61 (IRL, Oxford, 1995).
- Papadannou, V. & Johnson, R. in *Gene Targeting: a Practical Approach* (ed. Joyner, A. L.) 107–146 (IRL, Oxford, 1995).
- Gossler, A. & Zachgo, J. in *Gene Targeting: a Practical Approach* (ed. Joyner, A. L.) 181–225 (IRL, Oxford, 1995).
- Butcher, L. L. & Woolf, N. J. in *Handbook of Chemical Neuroanatomy* Vol II (eds Bjorkland, A., Hokfelt, T. & Kuhar, M. J.) Elsevier, Amsterdam, 1984).

ACKNOWLEDGEMENTS. M.P.S. and I.S.-S. contributed equally to this work. We thank S. Bryant, M. Garber, J. Guillo-Brown, L. K. Long, M. Swerdel and J. Wolf for technical assistance; T. Gridley for the C17 ES cells; J. Rossant for the SDKlacZpA plasmid; and L. Reichardt for the anti-p75 antibodies. J.F. was supported by the Wenner-Gren foundations.

CORRESPONDENCE and requests for materials should be addressed to M.B. (e-mail: barbacid@BMS.com).

Defects in enteric innervation and kidney development in mice lacking GDNF

José G. Pichel*, Liya Shen*, Hui Z. Sheng*, Ann-Charlotte Granholm†, John Drago‡, Alexander Grinberg*, Eric J. Lee*, Sing Ping Huang*, Mart Saarma§, Barry J. Hoffer†, Hannu Sariola§|| & Heiner Westphal*

* Laboratory of Mammalian Genes and Development, National Institutes of Health, Bethesda, 20892, USA

† Departments of Pharmacology and Basic Dental Science, University of Colorado Health Sciences Center, Denver, Colorado 80262, USA

‡ Department of Anatomy, Monash University, Clayton, 3168 Victoria, Australia

§ Programs of Molecular Neurobiology and Developmental Biology, Institute of Biotechnology, Biocentre 1, University of Helsinki, Viikinkaari 9, P.O. Box 56, FIN-00014 Helsinki, Finland

|| Laboratory of Pediatric Pathology, Childrens Hospital, University of Helsinki, Stenbäckinkatu 11, FIN-00380 Helsinki, Finland

GLIAL-CELL-LINE-DERIVED neurotrophic factor (GDNF) has been isolated as a neurotrophic factor for midbrain dopaminergic neurons¹. Because of its neurotrophic activity on a wide range of neuronal populations *in vitro* and *in vivo*^{2–11}, GDNF is being considered as a potential therapeutic agent for neuronal disorders^{2,3,12}. During mammalian development, it is expressed not only in the nervous system, but also very prominently in the metanephric kidney and the gastrointestinal tract, suggesting possible functions during organogenesis^{7,11,13,14}. We have investigated the role of GDNF during development by generating a null mutation in the murine *GDNF* locus, and found that mutant mice show kidney agenesis or dysgenesis and defective enteric innervation. We demonstrate that GDNF induces ureter bud formation

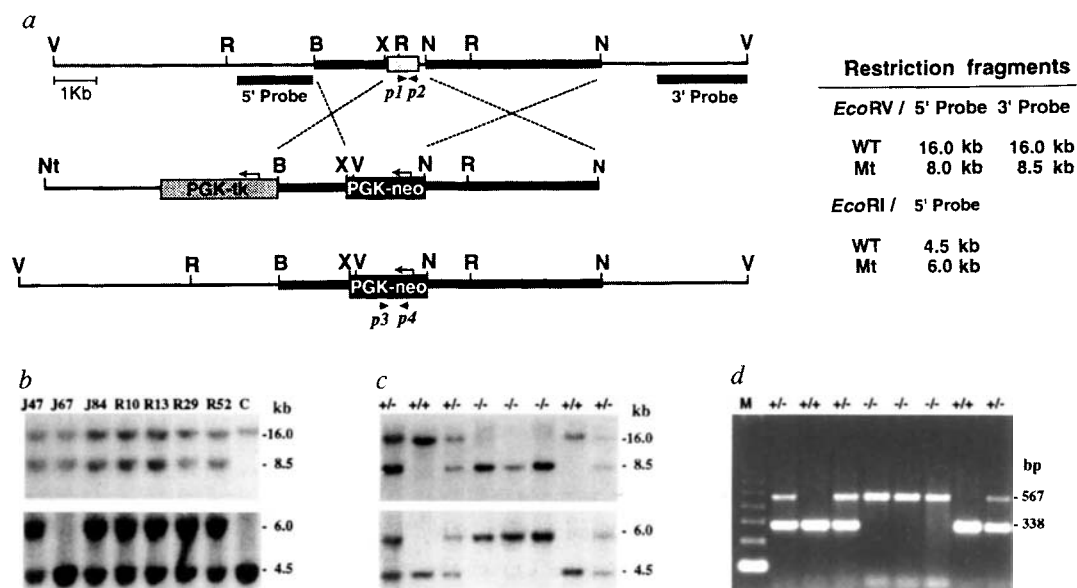
and branching during metanephros development, and is essential for proper innervation of the gastrointestinal tract.

We generated a non-functional allele of the *GDNF* gene by replacing part of the third exon that encodes GDNF protein^{1,13} with a cassette expressing the selectable marker neomycin phosphotransferase (*neo*) (Fig. 1a). After introducing this construct into embryonic stem (ES) cells, six clones were identified with the predicted mutant allele (Fig. 1b). Four clones produced chimaeric mice that transmitted the mutation to their progeny (Fig. 1c, d). Heterozygous offspring were viable and fertile, whereas mice homozygous for the mutant allele (*GDNF*^{−/−}) (Fig. 1c, d), although born in the expected mendelian frequency, died 12–24 h after birth. Dissection of *GDNF*^{−/−} mutant mice revealed either renal agenesis or severe renal dysgenesis (Fig. 2c). Surprisingly, heterozygous *GDNF*^{+/−} mice also showed renal abnormalities (Fig. 2b). When analysed macroscopically, 35% (*n* = 89) of the adult heterozygous mice between 3 and 5 months of age possessed a wide range of renal defects, including unilateral small kidneys with abnormal shape and/or cortical cysts or rough surface (23%), unilateral kidney agenesis, and severe bilateral kidney dysgenesis (12%). All other components of the urogenital system appeared macroscopically normal. The second abnormality found in *GDNF*^{−/−} mutant newborns was the presence of milk in the oesophagus and reduced progression of food into the gastrointestinal tract, suggesting a defect in peristalsis. These two major abnormalities correlate well with the reported patterns of *GDNF* expression during development, and suggest that GDNF functions in kidney and gastrointestinal development.

Metanephric kidney development is dependent on reciprocal inductive interactions between the ureteric bud, an outgrowth of the wolffian duct, and the adjacent metanephric mesenchyme¹⁵. Mesenchymal cells induce the ureteric bud to grow and branch repeatedly, forming the renal collecting system, and the ureteric bud induces the mesenchyme to condense into epithelial structures which differentiate into the glomeruli and the proximal and distal tubules of the nephron¹⁶. *GDNF* is highly expressed in the metanephric mesenchyme, in those cells that are induced to differentiate around the tip of the ureteric bud, although expression is downregulated by the epithelial differentiation of secretory nephrons^{13,14}. Histological analysis of *GDNF*^{−/−} embryos between embryonic day 11 (E11) and E14.5 showed that metanephric development is halted even though a morphologically distinct metanephric mesenchyme is present (Fig. 3c). Mesenchyme identity was confirmed by the presence of *Pax-2* signal at E12.5 even if no ureteric buds had formed (data not shown). No ureteric bud was found in most of the embryos (Fig. 3c). Where present, it had penetrated deeply into the mesenchyme, but there were fewer collecting ducts, nephric condensates and secretory nephrons (data not shown). Both types of defects, although less pronounced, were also seen in some of the heterozygous *GDNF*^{+/−} embryos (Fig. 3b).

Analysis of the branching pattern in *ex vivo* cultured urogenital blocks dissected from E11.5 embryos indicated a significant reduction in the growing arborization of the renal collecting system of heterozygotes (Fig. 3e), and a delayed or total absence of ureteric branching in the *GDNF*^{−/−} mutants (Fig. 3f), when compared with control explants (Fig. 3d). Differences in ureteric branching between the *GDNF*^{+/+}, *GDNF*^{+/−} and *GDNF*^{−/−} tissues was statistically significant (Fig. 3i). To examine whether GDNF was able to rescue the branching defect in mutant kidneys, we incubated explanted urogenital tissue in the presence of recombinant GDNF embedded in agarose beads. *In vitro* branching of the *GDNF*^{−/−} ureters was increased significantly (Fig 3h, i). Ectopic induction of supernumerary primary ureteric buds along the wolffian duct was occasionally observed in control *GDNF*^{+/+} samples incubated in the presence of the GDNF beads (Fig. 3g). Taken together, these results show that GDNF is a metanephric mesenchyme-derived factor, acting in a dose-dependent manner on the growth or arborization of the ureteric bud during early stages of kidney morphogenesis.

FIG. 1 Targeting the murine *GDNF* locus. **a**, Schematic representation of the *GDNF* locus (top), *GDNF* targeting vector (middle), expected mutant allele after homologous recombination (bottom), and restriction fragment lengths using the 5' or 3' probes (right). Successful gene replacement results in the deletion just under 1 kb of *GDNF* genomic sequence, including all of exon 3-derived coding sequences (grey box) (amino acids 52 to the C terminus)^{1,13}. In its place, a *PGK-neo* selection cassette²⁴ was inserted. Restriction enzyme sites: V, *EcoRV*; R, *EcoRI*; B, *Bam*HI; X, *Xba*I; N, *Nco*I; Nt, *Not*I. Symbols p1, p2, p3 and p4 indicate polymerase chain reaction (PCR) primers. WT, wild type; Mt, mutant. **b**, Southern blot analysis of targeted (J47, J67, J84, R10, R13, R29, R52) and parental (C) ES cell clones. Top, *EcoRV*-digested DNA when probed with the 3' flanking probe. Bottom, *EcoRI*-digested DNA hybridized with the 5' flanking probe. Clone J67 did not show the expected 5'-end homologous recombination. **c**, Southern blot analysis of progeny from a *GDNF*^{-/-} heterozygous intercross showing the three possible genotypes: wild type (+/+), heterozygous mutant (+/-), and homozygous mutant (-/-). Restriction enzymes and probes were as **b**. **d**, PCR assay used to genotype the *GDNF* locus. The 567-base-pair (bp) amplification product is generated from the mutant allele (neo primers p3 and p4), whereas the 338-bp



product (primers p1 and p2 within the 1.0-kb deleted segment) corresponds to the wild-type allele. DNA samples and distribution in lines shown are as **c**. M, molecular marker 123-bp ladder (GIBCO/BRL). **METHODS.** The targeting construct was linearized and transfected into J1 (ref. 25) or R1 (ref. 26) ES cell lines by electroporation. After G418 and gancyclovir selection²⁷ and screening by Southern analysis, 6 clones yielded the expected restriction fragments. Chimaeric mice were produced either by blastocyst injection²⁵ or by morula aggregation²⁶ using CD1 or C57BL/6 recipient strains to obtain germline transmission of the targeted allele.

The second phenotypic observation, an apparent defect in peristalsis in *GDNF*^{-/-} mutant mice, raised the possibility that enteric innervation might be affected by GDNF. GDNF transcripts have been detected at high levels in the outer mesenchymal layer throughout the length of the developing gastrointestinal tract^{13,14} known to be innervated by the neural crest-derived enteric nervous system (ENS)^{7,18}. Histological analysis of E18 embryos showed that the neurons of the myenteric plexus, readily apparent in normal embryos, were absent from the intestine of mutant *GDNF*^{-/-} mice (Fig. 4a, b). These observations were confirmed by immunofluorescence studies using antibodies against the neural markers choline acetyltransferase (ChAT) and neurofilament. A large number of enteric parasympathetic cholinergic ganglion cells were found distributed throughout the intestine of the normal *GDNF*^{+/+} mice (Fig. 4c) or heterozygous *GDNF*^{+/-} mutant embryos (data not shown); in contrast, the *GDNF*^{-/-} mutants were virtually devoid of these neurons (Fig. 4d). These data indicate a second major role for GDNF as a factor required for the proper innervation of the digestive tract. Whether GDNF is acting on multipotent neural crest-derived enteric ganglioblasts, or only on their survival after differentiation, is not yet known.

Based on the reported neurotrophic action of GDNF, notably in dopaminergic and cholinergic neuron populations, we examined several brain areas in the *GDNF* mutants. Our preliminary analysis failed to reveal overt differences in distribution, density or tyrosine hydroxylase staining properties between substantia nigra dopamine neurons in newborn *GDNF*^{-/-} mutant mice and the *GDNF*^{+/+} controls. The same was true for brain-stem motor nuclei and cortical cholinergic neurons evaluated by ChAT immunostaining. Thus no marked developmental disturbances of these neurons could be found. It is possible that GDNF-like molecules can compensate for the lack of GDNF function in the mutant brain. Because our mutant mice die shortly after birth, we are unable to assess a possible role for GDNF in the maintenance of these neurons after birth.

Several observations provide a possible link between GDNF function and that of the tyrosine kinase receptor c-Ret. This receptor is expressed in the budding ureter and in enteric ganglioblasts¹⁹, and *c-ret*^{-/-} mice manifest both renal agenesis or severe dysgenesis and lack of enteric neurons^{20,21}. GDNF may thus be involved in c-Ret-mediated signalling during ENS development and kidney organogenesis.

Some human genetic disorders, notably Hirschsprung disease

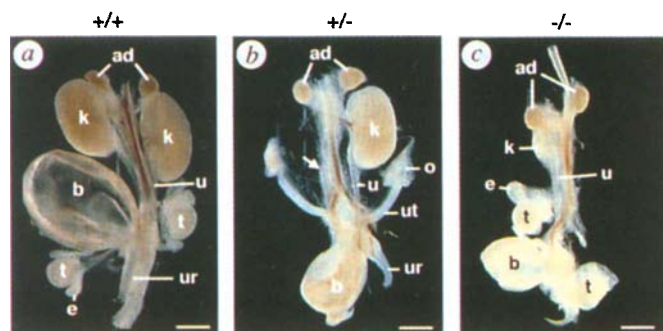


FIG. 2 Analysis of urogenital system defects in *GDNF* newborn mutant mice. **a**, Whole mount of a dissected urogenital system from a normal newborn male mouse. **b**, Right renal agenesis in a heterozygous *GDNF*^{+/-} female; the affected side reveals a blind-ending ureter without renal tissue at the tip (arrow). **c**, Renal agenesis and severe dysgenesis observed in a *GDNF*^{-/-} male. A rudimentary kidney with its ipsilateral ureter is present; the contralateral side shows neither kidney nor ureter. Different combinations of agenesis/dysgenesis are observed in mice with mutant genotypes. Adrenal glands, gonads and the rest of the urogenital system components appeared macroscopically normal. Abbreviations: ad, adrenal gland; b, bladder; e, epididymis; k, kidney; o, ovary; t, testes; u, ureter; ur, urethra; ut, uterus. Scale bars, 1 mm.

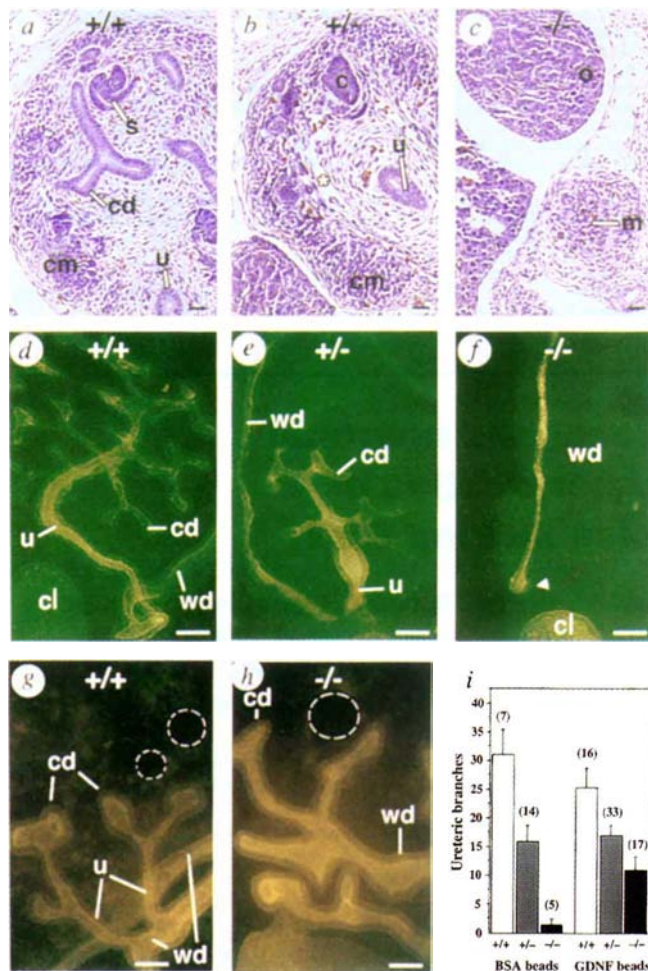
(HSCR) which causes megacolon formation, are characterized by defects in gastrointestinal innervation. Mutations in the human *RET* locus have been demonstrated in some familial forms of HSCR²². The loss of enteric neurons in *GDNF*^{-/-} mice suggest the possibility that mutations in the *GDNF* gene may also cause HSCR. Further, certain renal dysplasias that seem to result

from a defect of ureteric branching²³ might also be linked to *GDNF* gene defects.

In summary, *in vivo* characterization of the *GDNF* mutant phenotype and *ex vivo* explant culture of mutant tissues indicate that *GDNF* is essential for ureteric budding and branching, and the innervation of the gastrointestinal tract. It is possible that the

FIG. 3. Comparative analysis of the metanephros of *GDNF*^{+/+}, *GDNF*^{+/-} and *GDNF*^{-/-} embryos showing a ureteric branching defect in the mutant genotypes and its rescue by *GDNF* in *ex vivo* kidney cultures. *a-c*, Haematoxylin and eosin stained transverse sections of metanephros at E13.5. *a*, Section of a normal (*GDNF*^{+/+}) kidney; several branches of collecting ducts (cd) with condensing mesenchyme (cm) or s-shaped bodies (s) are clearly observed. *b*, In the *GDNF*^{-/-} mutant kidney, the ureter enters the kidney inducing condensation (cm) and epithelization (c) of mesenchyme, but there is reduced branching. The expansion of calyx-like spaces (asterisk) appears more enlarged along the entering ureter. *c*, The morphologically presumptive metanephric mesenchyme (m) of the *GDNF*^{-/-} mutant embryo in the absence of the ureteric bud remains histologically undifferentiated. *d-f*, Branching defect in mutant kidneys explanted at E11.5, incubated *in vitro* for 72 h and stained with anti-cytokeratin antibodies. The number of branches of the ureter is reduced in the *GDNF*^{-/-} animal (*e*) compared with the normal phenotype (*d*). *f*, A *GDNF*^{-/-} sample shows the wolffian duct with complete absence of ureteric budding. The tip of the wolffian duct (arrowhead) has not connected with the cloaca (cl). *g, h*, Recombinant *GDNF* action on ureteric branching in kidney explants. Addition of *GDNF*-containing agarose beads (broken circles) next to the explanted metanephric mesenchymes increases the branching of the wolffian duct and ureteric bud. *g*, An example of supernumerary ureteric budding in a normal kidney explant, showing two ureters that have budded from the wolffian duct and branched towards the bead. *h*, The branching towards the *GDNF* source is also seen in the 'rescued' *GDNF*^{-/-} specimens. Abbreviations: c, comma-shaped body; cd, collecting duct; cl, cloaca; cm, condensing mesenchyme; m, metanephric mesenchyme; s, s-shaped body; u, ureter; wd, wolffian duct. Scale bars: *a-c*, 50 μ m; *d-f*, 100 μ m; *g-h*, 50 μ m. *i*, Quantitative assessment of ureteric branching mediated by recombinant *GDNF*. Shown are numbers of collecting duct tips observed in *GDNF*^{+/+} (+/+), *GDNF*^{+/-} (+/-) and *GDNF*^{-/-} (-/-) urogenital specimens treated in culture in the presence of control BSA-soaked beads or *GDNF*-soaked beads. Values are means $\pm$ s.e.m. The difference in the number of branching tips between (+/+), (+/-) and (-/-) is statistically significant when using control BSA beads ($P < 0.01$, Student's *t*-test). Administration of *GDNF* in agarose beads significantly increases ureteric branching in homozygous *GDNF*^{-/-} samples (*GDNF* beads -/-) as compared to control BSA soaked beads (BSA beads -/-) ($P < 0.05$, paired *t*-test). Numbers in parenthesis indicate the sample size for each experimental condition. HCG/SF or TGF- β 1-soaked beads at the same concentration did not significantly effect the number of branches in the mutant specimens.

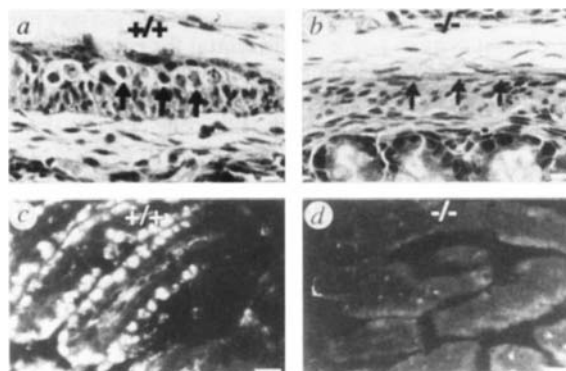
METHODS. E11.5 urogenital specimens were cultured on a 1.0- μ m Nucleopore filter (Costar) for 72 h in chemically defined Improved MEM Zinc Option medium (Richter's modification) (GIBCO) supplemented with 50 μ g ml⁻¹ transferrin (Sigma), fixed in cold methanol and whole-mount



immunostained as previously described²⁸. Affi-Gel blue agarose beads (Bio-Rad) saturated in recombinant *GDNF* or control factors at 100 ng μ l⁻¹ or in 1% BSA were placed on top of or next to the metanephric mesenchyme²⁹, and were incubated and analysed as described above.

FIG. 4. Histological and immunocytochemical analysis of normal and *GDNF*^{-/-} mutant gastrointestinal tracts. Sections through the gastrointestinal tract of *GDNF*^{+/+} (*a, c*) and homozygous *GDNF*^{-/-} mice (*b, d*). *a, b*, Haematoxylin and eosin staining of E18 embryos. The arrows in *a* point to neurons of the myenteric plexus that are identified by their large size, clear cytoplasm and prominent nuclei. These cells are absent in *b*. Arrows point to their expected position. *c, d*, Density and localization of enteric ganglia in neonatal mice using antibodies directed against choline acetyltransferase (ChAT). Many ChAT-positive neurons are present in the intestinal wall of normal mice. None are seen in the *GDNF*^{-/-} mice. Scale bars: *a, b*, 25 μ m; *c, d*, 32 μ m.

METHODS. Intestines from *GDNF*^{+/+} and *GDNF*^{-/-} and normal mice were immersion fixed in 4% paraformaldehyde in phosphate-buffered saline (PBS) overnight and stored in 30% sucrose in PBS. After embedding in OCT medium (Miles), 20- μ m cryostat coronal sections were collected and mounted on glass slides. Immunostaining was performed as described³⁰ using antibodies directed against ChAT (Boehringer-Mannheim). FITC-conjugated anti-rat IgG (Dako) was used as secondary antibody.



GDNF locus is implicated in human disorders in which target fields of this factor are affected.

Note added in proof: Results similar to ours are reported by Sánchez *et al.* and by Moore *et al.* (*Nature* **381**, 789–792, 1996) and Treanor *et al.* (this issue) provide direct evidence that *GDNF* is involved in cRet-mediated signalling. □

Received 9 April; accepted 5 June 1996.

- Lin, L.-F. H., Doherty, D. H., Lile, J. D., Bektess, S. & Collins, F. *Science* **260**, 1130–1132 (1993).
- Tomac, A. *et al.* *Nature* **373**, 335–339 (1995).
- Beck, K. D. *et al.* *Nature* **373**, 339–341 (1995).
- Henderson, C. E. *et al.* *Science* **266**, 1062–1064 (1994).
- Yan, Q., Matheson, C. & Lopez, O. T. *Nature* **373**, 341–344 (1995).
- Oppenheim, R. W. *et al.* *Nature* **373**, 344–346 (1995).
- Trupp, M. *et al.* *J. Cell Biol.* **130**, 137–148 (1995).
- Buj-Bello, A., Buchman, V. L., Horton, A., Rosenthal, A. & Davies, A. M. *Neuron* **15**, 821–828 (1995).
- Arenas, E., Trupp, M., Åkerud, P. & Ibáñez, C. *Neuron* **15**, 1465–1473 (1995).
- Mount, H. T. J., Dean, D. O., Alberch, J., Dreyfus, C. F. & Black, I. B. *Proc. natn. Acad. Sci. U.S.A.* **92**, 9092–9096 (1995).
- Poulsen, K. T. *et al.* *Neuron* **13**, 1245–1252 (1994).
- Gash, D. M. *et al.* *Nature* **380**, 252–255 (1996).
- Heilmich, H. L., Kos, L., Cho, E. S., Mahon, K. A. & Zimmer, A. *Mech. Dev.* **54**, 95–105 (1996).
- Suvanto, P. *et al.* *Eur. J. Neurosci.* **8**, 816–822 (1996).

- Saxén, L. *Organogenesis of the Kidney* (Cambridge Univ. Press, 1987).
- Eklom, P. *FASEB J.* **3**, 2141–2150 (1989).
- Yntema, C. L. & Hammond, W. S. *J. comp. Neurol.* **101**, 515–542 (1954).
- Le Douarin, N. M. & Teillet, M.-A. *J. Embryol. exp. Morphol.* **30**, 31–48 (1973).
- Pachnis, V., Mankoo, B. & Costantini, F. *Development* **119**, 1005–1017 (1993).
- Schuchardt, A., D'Agati, V., Larsson-Blomberg, L., Costantini, F. & Pachnis, V. *Nature* **367**, 380–383 (1994).
- Durbec, P. L., Larsson-Blomberg, L. B., Schuchardt, A., Costantini, F. & Pachnis, V. *Development* **122**, 349–358 (1996).
- Pasini, B., Ceccherini, I. & Romeo, G. *Trends Genet.* **12**, 138–144 (1996).
- Osathanondh, V. & Potter, E. L. *Archs Path.* **82**, 391–402 (1966).
- Tybuliewicz, V. L. J., Crawford, C. E., Jackson, P. K., Bronson, R. T. & Mulligan, R. C. *Cell* **65**, 1153–1163 (1991).
- Li, E., Bestor, T. H. & Jaenisch, R. *Cell* **69**, 915–926 (1992).
- Nagy, A., Rossant, J., Nagy, R., Abramow-Newerly, W. & Roder, J. C. *Proc. natn. Acad. Sci. U.S.A.* **90**, 8424–8428 (1993).
- Love, P. E., Tremblay, M. & Westphal, H. *Proc. natn. Acad. Sci. U.S.A.* **89**, 9929–9933 (1992).
- Sariola, H., Holm, K. & Henke-Fahle, S. *Development* **104**, 589–599 (1988).
- Vainio, S., Karavanova, I., Jowett, A. & Thesleff, I. *Cell* **75**, 45–58 (1993).
- Harlow, E. & Lane, D. *Antibodies. A Laboratory Manual* (Cold Spring Harbour Laboratory Press, NY, 1988).

ACKNOWLEDGEMENTS. We thank D. F. Porter and D. Feltner for comments on drafts of this manuscript. B.J.H. thanks N. Srivastava and J. Mott for assistance with histological evaluation of the nervous tissues. J.D. is a Bassett fellow of the Royal Australasian College of Physicians, supported by the NH and MRC. M.S. and H.S. were supported by a grant from the Academy of Finland. B.J.H. received a grant from the NIH.

CORRESPONDENCE. and requests for materials should be addressed to H.W. (e-mail: hw@helix.nih.gov). Details of the vector and PCR procedure used in Fig. 1 are available on request.

Renal and neuronal abnormalities in mice lacking *GDNF*

Mark W. Moore*, Robert D. Klein†, Isabel Fariñas‡, Hansjorg Sauer†, Mark Armanini†, Heidi Phillips†, Louis F. Reichardt‡, Anne M. Ryan§, Karen Carver-Moore* & Arnon Rosenthal†

Departments of *Molecular Biology, †Neuroscience, and §Pathobiology, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA

‡Department of Physiology and Howard Hughes Medical Institute, University of California, Room U-332, Third and Parnassus Avenue, San Francisco, California 94143, USA

GLIAL cell-line derived neurotrophic factor (*GDNF*) is a potent survival factor for embryonic midbrain dopaminergic¹, spinal motor², cranial sensory³, sympathetic, and hindbrain noradrenergic⁴ neurons, and is available to these cells *in vivo*. It is therefore considered a physiological trophic factor and a potential therapeutic agent for Parkinson's disease^{5,6}, amyotrophic lateral sclerosis⁷, and Alzheimer's disease⁸. Here we show that at postnatal day 0 (P0), *GDNF*-deficient mice have deficits in dorsal root ganglion, sympathetic and nodose neurons, but not in hindbrain noradrenergic or midbrain dopaminergic neurons. These mice completely lack the enteric nervous system (ENS), ureters and kidneys. Thus *GDNF* is important for the development and/or survival of enteric, sympathetic and sensory neurons and the renal system, but is not essential for catecholaminergic neurons in the central nervous system (CNS).

The murine *GDNF* gene was disrupted by homologous recombination in embryonic stem (ES) cells (Fig. 1*a, b*), and a targeted clone was injected into blastocysts to generate *GDNF*-mutant mice (Fig. 1*c*). Whereas *GDNF* messenger RNA was found in the kidney, intestine, ventral midbrain and skeletal muscle of normal embryonic day 13.5 (E13.5) mice, no *GDNF* transcripts could be detected in littermates homozygous for the mutant allele (*GDNF*^{−/−}) (data not shown). Heterozygous mice were normal in size and were indistinguishable, by visual inspection of live animals, from their wild-type littermates. In contrast, although *GDNF*^{−/−} mice were able to suckle and had normal limb and body movements, they died 1–1.5 days after birth.

GDNF, which is expressed in the embryonic striatum (a major innervation target for dopaminergic neurons)^{8–10}, was originally characterized by its ability to prevent the death of embryonic midbrain dopaminergic neurons in culture¹ and in lesion models *in vivo*^{11–14}. We therefore examined whether *GDNF* is an essential survival factor for dopaminergic neurons during normal development. Surprisingly, the number of tyrosine hydroxylase positive (TH⁺) dopaminergic neurons in the substantia nigra (Table 1 and Fig. 1*d, e*) and the apparent density of dopaminergic projections to the striatum (Fig. 1*f, g*) were identical in *GDNF*^{−/−} and wild-type mice. No statistically significant deficit in the number of dopaminergic fibres (data not shown) were detected in postnatal day 42 (P42) *GDNF*^{+/−} heterozygous mice. These results demonstrate that *GDNF* is not a required survival or neurite-outgrowth factor for embryonic dopaminergic neurons, and suggest that it is not a limiting survival factor for dopaminergic neurons, which continue to mature postnatally¹⁵, as had been previously assumed^{1,11–14}.

After its isolation, *GDNF* was shown to be a potent neurotrophic factor for embryonic spinal motor neurons in culture² and *in vivo*^{16,17}. We therefore determined whether *GDNF*, which is expressed in skeletal muscle, is required for survival of motor neurons. No deficits in the facial motor neurons of P0 *GDNF*^{−/−} or of P42 *GDNF*^{+/−} animals were observed (Table 1), but a small (yet statistically significant) loss of motor neurons was detected in the lumbar spinal (22%) and trigeminal (19%) ganglia of P0 *GDNF*^{−/−} mice (Table 1 and Fig. 1*h, i*). These findings indicate that *GDNF* is not a critical neurotrophic factor for the majority of voluntary motor neurons during embryogenesis.

GDNF was recently shown to prevent the chemically induced death of noradrenergic neurons in the locus coeruleus, and to promote their fasciculation and sprouting in whole animals⁴. Noradrenergic neurons in the locus coeruleus were found to be approximately normal in size (data not shown) and in number in

FIG. 1 Disruption of the *GDNF* gene and the resulting neuronal phenotype. **a**, Targeting construct, wild-type *GDNF* allele, and the disrupted allele. Amino acids 103–211 of the mature biologically active portion of *GDNF* are missing from the disrupted allele. The location of the probe used in the Southern blot is indicated. The directions of gene transcription are marked by horizontal arrows. **b**, Detection of homologous recombination event in an ES clone by Southern hybridization. **c**, Genotype analysis of wild-type (+/+), heterozygous (+/−) and homozygous mutant (−/−) animals by polymerase chain reaction (PCR). The upper band is specific for the *neo*^r gene, and the lower band is specific for the wild-type *GDNF* gene. **d–o**,

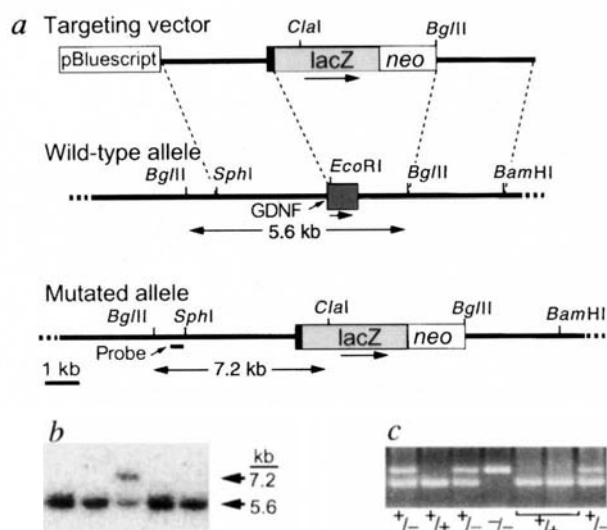
TABLE 1 Neuronal counts

Neuron type	Wild-type mice	N	No. of -/- or +/- mice	N	Reduction (%)
(a) P0 wild-type and <i>GDNF</i> ^{-/-} mice					
CNS catecholaminergic neurons					
Dopaminergic (SN)	4,242 ± 114	2	4,347 ± 261	2	n.s.
Noradrenergic (LC)	2,016 ± 48	2	1,704 ± 30	2	n.s.
Motor nuclei					
Facial (VII)	3,453 ± 60	3	3,264 ± 198	5	n.s.
Trigeminal (V)	1,152 ± 60	3	928 ± 71	4	19*
Spinal lumbar (LI-L6)	3,488 ± 206	4	2,734 ± 138	3	22**
Sensory ganglia					
Trigeminal	42,144 ± 1,707	3	37,275 ± 3,252	3	n.s.
Vestibular	3,495 ± 189	3	3,658 ± 562	3	n.s.
Petrosal + nodose	9,666 ± 271	4	5,764 ± 194	4	40***†
L5 dorsal root	10,937 ± 367	4	8,448 ± 358	4	23***†
Sympathetic ganglion					
Superior cervical	29,949 ± 3,329	4	19,319 ± 772	4	35*†
(b) P42 wild-type and <i>GDNF</i> ^{-/-} mice					
Dopaminergic (SN)	118.04 ± 7.34	8	112.88 ± 9.04	12	n.s.
Noradrenergic (LC)	1,218.5 ± 91.24	8	1,068 ± 38.19	12	n.s.
Facial motor (VII)	1,701 ± 55.8	8	1,657 ± 54.4	12	n.s.

Values are expressed as mean number of neurons ± s.e.m. *N*, number of mice analysed; n.s., not statistically significant; one-tailed Student's *t*-test; SN, substantia nigra; LC, locus coeruleus; L5, fifth lumbar. (a) Analysis of deficits in P0 wild-type and homozygous mutant mice. CNS catecholaminergic neurons were quantified in TH-stained paraffin sections. Analysis of 3 additional *GDNF*^{-/-} and 5 wild-type mice using a different counting procedure¹⁴ (not shown) confirmed the lack of deficits. PNS and motor neurons were counted in cresyl violet-stained paraffin sections 7 µm thick. Neurons were counted in every sixth section. (b) Analysis of P42 wild-type and heterozygous animals. Quantification of TH-immunoreactive cell numbers in the pars compacta region of the substantia nigra was conducted as described²⁴. Cell counts are represented as the mean number of cells per section per animal. Cell counts for the LC (TH-positive) and facial nuclei (cresyl violet stained) were derived from analysis of a series of 30-µm sections. Every third section was counted on both sides of the brain stem. Cell numbers represent mean cumulative raw counts per animal.

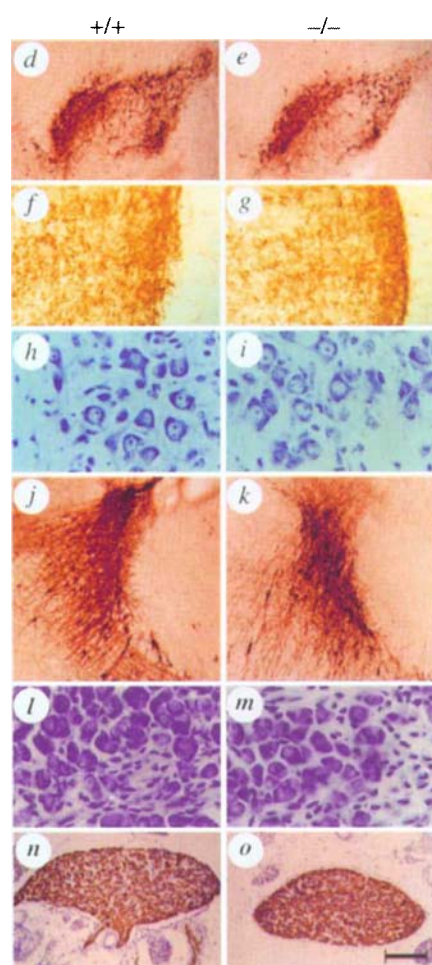
*** *P* < 0.01; ** *P* < 0.025; * *P* < 0.05.

† Because changes in cell size were detected in these populations (see text), cell numbers were also determined using the optical dissector²⁹ method and similar deficits were found. Numbers obtained by this method were (+/+, -/-, % reduction): petrosal-nodose, 22,937 ± 2,398, 13,758 ± 5,192, 40%; L5 dorsal root ganglion, 17,234 ± 2,608, 12,415 ± 1,537, 28%; superior cervical ganglion, 70,560 ± 8,241, 42,171 ± 3,006, 40%.



Panels compare neuronal populations in P0 wild-type (+/+) and *GDNF*^{-/-} (-/-) mice. Tyrosine hydroxylase (TH) staining of: d, e, substantia nigra; f, g, striatum; j, k, locus coeruleus; n, o, superior cervical ganglion. TH is the rate-limiting enzyme in dopamine synthesis. Cresyl violet staining of spinal motoneurons (h, i) and petrosal-nodose ganglia (l, m) neurons.

METHODS. A 3-kb *SphI*-*EcoRI* genomic fragment encoding amino acids 52-102 was fused in-frame to the *lacZ* gene. A *neo* gene under the control of the phosphoglycerate kinase 1 promoter and a 3.1-kb *BglII*-*BamHI* fragment from the 3' end of the *GDNF* gene were inserted immediately downstream of the *lacZ* gene. The targeting construct was electroporated into ES-D3.12 cells, and 400 µg ml⁻¹ G418 was used to isolate resistant clones. Targeted clones were injected into C57BL/6 blastocysts and a single clone was transmitted through the germ line. Neuronal analysis was performed as described²⁰. TH antibodies from Pel-Freez were used at 1:200. Cell areas were measured with an image analysis system using the MCID-M4 v.1.2 program (Imaging Research, Ontario). Scale bar: d, e, 300 µm; f, g, 50 µm; h, i, 25 µm; j, k, 250 µm; l, m, 8 µm; n, o, 200 µm.



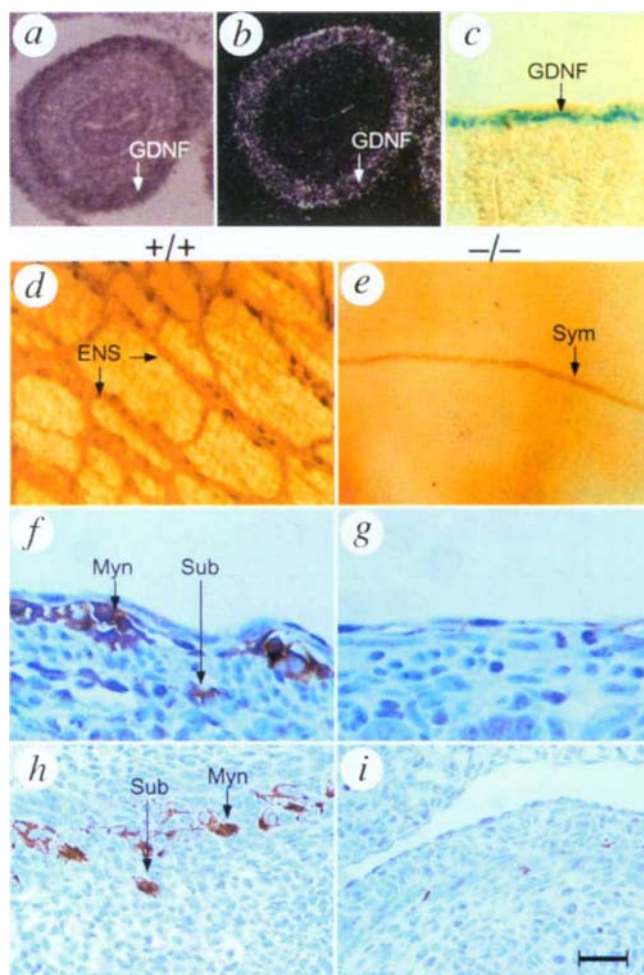


FIG. 2 ENS in wild-type and $GDNF^{-/-}$ mice. *a, b*, Expression of $GDNF$ mRNA in E13.5 mouse gut as detected by *in situ* hybridization. *c*, Cross-section through the small intestine from a P0 $GDNF^{-/-}$ mouse stained for β -galactosidase demonstrates expression of the targeted gene in a subset of muscle cells, close to where the ENS is located. *d–i*, Whole mounts of small intestine from P0 wild-type (+/+) and $GDNF^{-/-}$ (-/-) mice stained with antibodies to NSE (*d, e*) and peripherin (*f, g*). Small intestine from E13.5 wild-type (*h*) and $GDNF^{-/-}$ (*i*) embryos stained with peripherin. Myenteric (Myn) and submucosal (Sub) neurons, shown by arrows in the wild type, are absent in $GDNF^{-/-}$ mice. (ENS neurons were not found in the intestine and very rarely found in the stomach of $GDNF^{-/-}$ mice.) ENS neurons are indicated. Sites of $GDNF$ mRNA expression are marked (GDNF). Afferent fibres, probably derived from sympathetic innervation to the gut, are also marked (Sym). Scale bars: *a, b*, 65 μ m; *d–i*, 5 μ m.

METHODS. Animals were perfusion fixed with 10% neutral buffered formalin, paraffin embedded, and sectioned at 5 μ m for light microscopic analysis. Intestines from animals perfused with 4% paraformaldehyde were stained in whole mount with antibodies to neural markers or with X-gal. Antibody staining was performed as described³⁰, using polyclonal anti-peripherin (1:300; Chemicon) or anti-NSE (1:10, Incstar). *In situ* hybridization to normal E13.5 mouse embryos was performed as described².

both the P0 $GDNF^{-/-}$ (Table 1 and Fig. 1j, k) and the P42 $GDNF^{+/+}$ mice (Table 1), indicating that GDNF is not an obligatory physiological survival factor for this neuronal population. Similarly, although GDNF is upregulated in the hippocampus, cortex and striatum after chemically induced epileptic seizure or injection of excitatory neurotransmitters^{18,19}, no gross deficits were identified in the cerebellum, basal forebrain, hippocampal formation, striatum or neocortex of P0 $GDNF^{-/-}$ mice (data not shown). These findings combined suggest that the presence of GDNF in the embryonic CNS and in innervation targets may reflect, at least in part, involvement in functions other than neuronal survival.

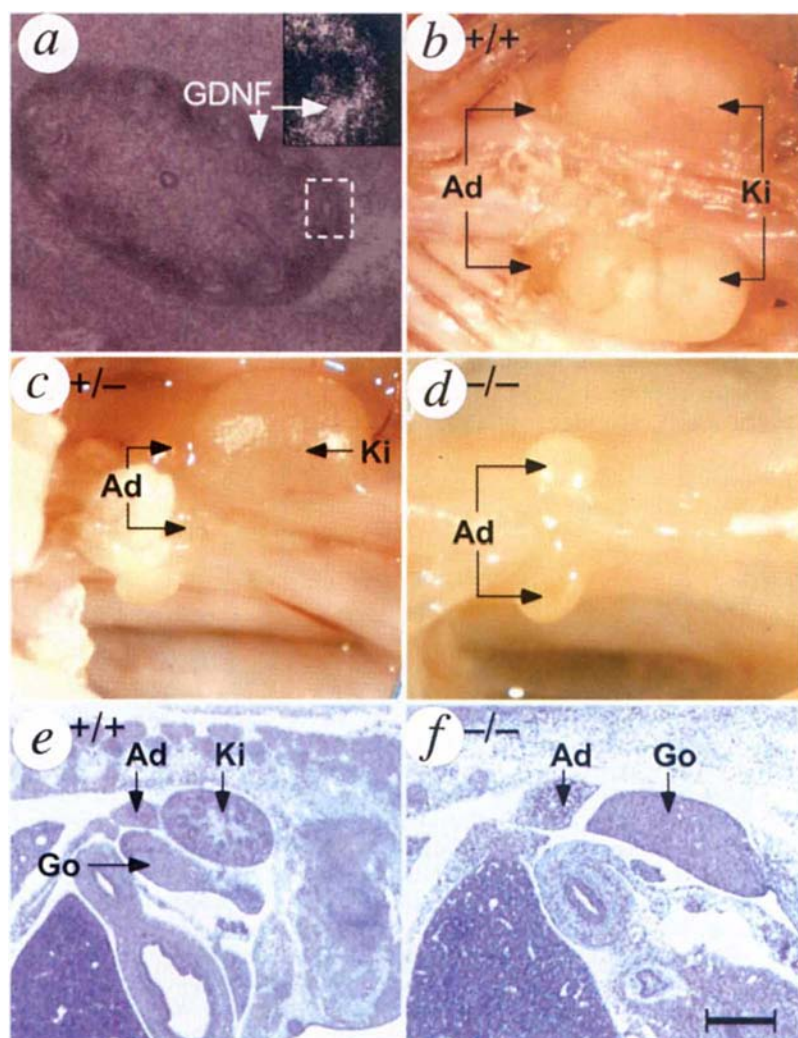
Consistent with the observation that GDNF promotes the survival of chick embryonic sympathetic and nodose neurons in culture³, there was a reduction in the number of neurons in petrosal–nodose (40%) (Fig. 1l, m), sympathetic superior cervical (SCG) (35%) (Fig. 1n, o), and dorsal root (23%) ganglia of P0 $GDNF^{-/-}$ mice (Table 1). In addition, the neurons in the affected ganglia were smaller than in the wild-type counterparts (cell areas were (-/-; +/+) $83.6 \pm 10.9 \mu\text{m}^2$ compared with $116 \pm 6.0 \mu\text{m}^2$, $49.0 \pm 2.7 \mu\text{m}^2$ compared with $64.5 \pm 6.2 \mu\text{m}^2$, and $81.1 \pm 10.5 \mu\text{m}^2$ compared with $125.3 \pm 14.7 \mu\text{m}^2$ for petrosal–nodose, sympathetic superior cervical and dorsal root ganglia, respectively), indicating that GDNF modulates the survival and possibly the maturation of these autonomic and sensory neurons. In contrast, no significant reduction in size or number of trigeminal sensory and vestibular ganglia neurons was detected (Table 1 and data not shown).

In situ hybridization analysis indicated abundant expression of $GDNF$ mRNA in the muscle layers of the intestine during embryogenesis²⁰ (Fig. 2a–c), so we examined myenteric (Auerbach) and submucosal (Meissner) plexi for possible neuronal deficits. ENS neurons belonging to these two plexi were readily visible along the length of the gastrointestinal tract of E12.5 to P0 wild-type and $GDNF^{+/+}$ mice in haematoxylin and eosin-stained sections (data not shown), or using antibodies to the neurofilament subunit of relative molecular mass 150,000 (M_r 150K) (data not shown), neuronal-specific enolase (Fig. 2d) or peripherin (Fig. 2f, h). In contrast, these vagal neural crest-derived neurons were completely absent in the small and large intestine of age-matched $GDNF^{-/-}$ littermates (Fig. 2e, g, i), and only a few neurons were detected in the stomach of these mice (data not shown). The absence of peripherin-positive cells from the gut of $GDNF^{-/-}$ mice as early as E12.5 suggests that GDNF may be essential for the commitment, migration or differentiation of ENS neuron precursors²¹. However, we cannot exclude the possibility that GDNF is an early survival factor in the ENS. In addition, the smooth muscles in the intestinal wall were thinner in $GDNF^{-/-}$ mice compared to wild-type mice. Absence of the ENS was previously noted in mice lacking the orphan tyrosine kinase receptor Ret^{22,23}.

In accordance with the presence of $GDNF$ mRNA in the embryonic kidney mesenchyme from mice aged E11.5 or older²⁰ (Fig. 3a), 16 of 22 $GDNF^{-/-}$ mice had complete bilateral renal and ureteral agenesis (Fig. 3d, f) (partial microscopic development of one kidney and ureter was observed in 6 $GDNF^{-/-}$ embryos). In heterozygous embryos, pups, and adults of both sexes, there was an increased incidence of unilateral renal agenesis (7 of 26) or hypoplasia (4 of 26) relative to wild-type mice (0 of 12) (Fig. 3b, c, e). Other derivatives of embryonic urogenital intermediate mesoderm (pronephros and mesonephros, adrenal gland and gonads), remaining abdominal viscera and all thoracic tissues were microscopically normal in both $GDNF^{+/+}$ and $GDNF^{-/-}$ mice, indicating that GDNF is required specifically for the development of the metanephric kidney. Examination of sections stained with haematoxylin and eosin throughout the kidney region of $GDNF^{-/-}$ mice at E11.5, when the metanephric kidney starts to form, demonstrated the presence of a mesonephric duct but not of a morphologically defined ureter or ureteric bud. In addition, no properly developed metanephric blastema could be detected in these mice (data not shown). A morphologically normal ureteric bud and metanephric blastema were present in one of four E11.5 $GDNF^{-/-}$ mice analysed (all four mice were serially sectioned at 5 μ m, and every section from the kidney region was examined). The metanephric kidney develops by reciprocal inductive interactions between the ureteric bud, an evagination of the mesonephric/wolffian duct, which gives rise to the collecting ducts/ureter, and the metanephric blastema, the caudal intermediate mesodermal precursor of the renal parenchyma, which gives rise to the renal tubules of the metanephric (definitive) kidney²⁴. The pattern of expression of GDNF in the metanephric blastema combined with the absence of a morphologically defined ureteric

FIG. 3 Kidneys in wild-type $GDNF^{+/+}$ and $GDNF^{-/-}$ mice. **a**, *In situ* hybridization of GDNF in E13.5 mouse kidney. As shown in the inset, GDNF is expressed in the undifferentiated mesenchyme but not in developing nephrons. **b–d**, Photographs of the abdomen in P0 wild-type (**b**), $GDNF^{+/-}$ (**c**) or $GDNF^{-/-}$ (**d**) mice. Note the position of the kidneys subadjacent to the adrenals in **b**, their absence in the mutant (**d**), and the absence of one kidney in a heterozygote (**c**). **e**, **f**, Haematoxylin and eosin-stained sagittal sections of E13.5 wild-type ($+/+$) and $GDNF^{-/-}$ ($-/-$) embryos. **f**, Gonads (Go) are present in the space normally occupied by the kidney (Ki), just caudal to the adrenal (Ad). Sites of expression of GDNF mRNA are indicated. Scale bar: **a**, 130 μ m; **b–d**, 500 μ m; **e**, **f**, 50 μ m.

METHODS. Wild-type $GDNF^{+/+}$ and $GDNF^{-/-}$ mice were killed at the indicated gestational age, fixed with 10% neutral buffered formalin, embedded in paraffin, serially sectioned, and stained with haematoxylin and eosin for microscopic examination.



bud raises the possibility that GDNF may participate in the induction of the ureteric bud by the mesonephric duct. Alternatively, GDNF may control the proliferation or survival of ureteric bud cells after they have been specified. In the absence of a ureteric bud the mesenchyme will not develop into renal tubules and no kidney will be formed. Finally, it is conceivable that GDNF participates in the commitment, survival, proliferation or differentiation of the metanephric blastema, and that a functional deficiency in the mesenchyme of the $GDNF^{-/-}$ mice leads to the absence of a ureteric bud. Defects in kidney development (as well as in other organs) were observed in mice deficient in bone morphogenic factor (BMP)-7 (ref. 25), Ret²² and the Wilms tumour-associated, putative transcription-inhibiting factor, WT-1 (ref. 26).

With respect to the reproductive organs, the single noted change in $GDNF^{-/-}$ mice was a reversal in the orientation of the ovary in relation to the abdominal viscera (data not shown).

Additionally, $GDNF^{-/-}$ mice had mild multifocal necrosis in the splenic red pulp at sites of active haematopoiesis (data not shown).

In summary, our findings demonstrate that GDNF is not a primary neurotrophic factor for dopaminergic motor or noradrenergic neurons during embryogenesis. Rather, GDNF seems to be essential for development of the ENS, kidney and subpopulations of sensory and sympathetic neurons. The striking similarity in phenotype between the GDNF-deficient mice described here and Ret-deficient mice^{22,23}, combined with the expression of Ret in all of the GDNF-responsive tissues^{27,28}, raises the possibility that these two proteins function in the same signalling cascade. □

Received 19 March; accepted 13 May 1996.

- Lin, L. H., Doherty, D. H., Lile, J. D., Bektess, S. & Collins, F. *Science* **260**, 1130–1132 (1993).
- Henderson, C. E. *et al. Science* **266**, 1062–1064 (1994).
- Bui-Bello, A., Buchman, V. L., Horton, A., Rosenthal, A. & Davies, A. M. *Neuron* **15**, 821–828 (1995).
- Arenas, E., Trupp, M., Akerud, P. & Ibáñez, C. F. *Neuron* **15**, 1465–1473 (1995).
- Hirsch, E. C., Graybiel, A. M. & Agid, Y. A. *Nature* **334**, 345–348 (1988).
- Hornykiewicz, O. *Mt Sinai J. Med.* **55**, 11–20 (1988).
- Hirano, A. in *Anyotrophic Lateral Sclerosis and Other Motor Neuron Diseases* (ed. Rowland, L. P.) 91–101 (Raven, New York, 1991).
- Schaar, D. G., Sieber, B. A., Dreyfus, C. F. & Black, I. B. *Exp. Neurol.* **124**, 368–371 (1993).
- Strömberg, I. *et al. Exp. Neurol.* **124**, 401–412 (1993).
- Poulsen, K. T. *Neuron* **13**, 1245–1252 (1994).
- Tomac, A. *et al. Nature* **373**, 335–339 (1995).
- Beck, K. *et al. Nature* **373**, 339–341 (1995).
- Keams, C. M. & Gash, D. M. *Brain Res.* **672**, 104–111 (1995).
- Sauer, H. *et al. Proc. natn. Acad. Sci. U.S.A.* **92**, 8935–8939 (1995).
- Coyle, J. T. & Campochiaro, P. J. *Neurochem.* **27**, 673–678 (1976).
- Yan, Q., Matheson, C. & Lopez, O. T. *Nature* **373**, 341–344 (1995).
- Oppenheim, R. *et al. Nature* **373**, 344–346 (1995).
- Schmidt-Kastner, R. *et al. Brain Res. molec. Brain Res.* **26**, 325–330 (1994).

- Humpel, C. *et al. Neuroscience* **59**, 791–795 (1994).
- Hellmich, H. L., Kos, L., Cho, E. S., Mahon, K. A. & Zimmer, A. *Mech. Dev.* **54**, 95–105 (1996).
- Gershon, M. D., Chalazonitis, A. & Rothman, T. P. *J. Neurobiol.* **2**, 199–214 (1993).
- Schuchardt, A., D'Agati, V., Larsson-Blomberg, L., Costantini, F. & Pachnis, V. *Nature* **367**, 380–383 (1994).
- Durbec, P. L., Larsson-Blomberg, L. B., Schuchardt, A., Costantini, F. & Pachnis, V. *Development* **122**, 349–358 (1996).
- Saxen, L. *Organogenesis of the Kidney* Vol. 19 (eds Barlow, P. W., Green, P. B. and White, C. C.) (Cambridge Univ. Press, 1987).
- Dudley, A. T., Lyons, K. M. & Robertson, E. J. *Genes Dev.* **9**, 2795–2807 (1995).
- Kreidberg, J. A. *et al. Cell* **74**, 679–691 (1993).
- Pachnis, V., Mankoo, B. & Costantini, F. *Development* **119**, 1005–1017 (1993).
- Liching, L. & Anderson, D. J. *Neuron* **15**, 527–539 (1995).
- Gundersen, H. J. G. *J. Microsc.* **143**, 3–45 (1986).
- Jones, K. R., Fariñas, I., Backus, C. & Reichardt, L. F. *Cell* **76**, 989–999 (1994).

ACKNOWLEDGEMENTS. M.W.M., R.D.K. and I.F. contributed equally to this study. We thank M. Dowd, M. Bauer and S. Pitts-Meek for help with the generation of the GDNF-deficient mice; E. Bery and W. Anstine for preparing the manuscript and figures; C. Kyle-Byrne for peripheral immunohistochemistry; and K. S. O'Shea for critical reading of the manuscript. I.F. is supported by the Human Frontier Science Program. L.F.R. is an investigator of the Howard Hughes Medical Institute.

CORRESPONDENCE and requests for materials should be addressed to A.R. (e-mail: ar@gene.com).

Characterization of a multicomponent receptor for GDNF

James J. S. Treanor*, Laurie Goodman*, Frederic de Sauvage†, Donna M. Stone*, Kris T. Poulsen*, Claus D. Beck*, Christa Gray‡, Mark P. Armanini*, Richard A. Pollock§, Franz Hefti*, Heidi S. Phillips*, Audry Goddard‡, Mark W. Moore‡, Anna Buj-Bello||, Alun M. Davies||, Naoya Asai¶, Masahide Takahashi¶, Richard Vandlen#, Christopher E. Henderson§ & Arnon Rosenthal*

Departments of * Neuroscience, ‡ Molecular Biology, # Protein Biochemistry, and † Molecular Cancer Research, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA
§ INSERM U.382, IBDM, Campus de Luminy—Case 907, 13288 Marseille, Cedex 09, France
|| University of St Andrews, Bute Medical Building, St Andrews, Fife KY16 9AT UK
¶ Department of Pathology, Nagoya University, Nagoya 466, Japan

GLIAL-CELL-LINE-DERIVED neurotrophic factor (GDNF)¹ is a potent survival factor for central and peripheral neurons^{2–6}, and is essential for the development of kidneys and the enteric nervous system^{7–9}. Despite the potential clinical and physiological importance of GDNF, its mechanism of action is unknown. Here we show that physiological responses to GDNF require the presence of a novel glycosyl-phosphatidylinositol (GPI)-linked protein (designated GDNFR- α) that is expressed on GDNF-responsive cells and binds GDNF with a high affinity. We further demonstrate that GDNF promotes the formation of a physical complex between GDNFR- α and the orphan tyrosine kinase receptor Ret^{10–12}, thereby inducing its tyrosine phosphorylation. These findings support the hypothesis that GDNF uses a multi-subunit receptor system in which GDNFR- α and Ret function as the ligand-binding and signalling components, respectively.

In searching for a receptor for GDNF, embryonic rat midbrain enriched for GDNF-responsive dopaminergic neurons was used to generate a complementary DNA library in a cytomegalovirus-based expression vector. The library was divided into 400 pools each containing 1,500 cDNA clones, and each pool was transfected into COS 7 cells and screened for the presence of GDNF-binding proteins¹³. One positive pool was identified and was repeatedly subdivided until a single cDNA clone, encoding a novel protein 468 amino acids long, was isolated (Fig. 1a). Analysis of this protein (designated GDNFR- α) indicated the presence of a signal peptide for secretion at the amino terminus, a stretch of 23 hydrophobic amino acids at the carboxy terminus, 3 glycosylation sites, and 30 cysteine residues which appear to have a similar topology to those found in many cytokine receptors¹⁴ (Fig. 1a). The hydrophobic cluster at the C terminus is preceded by a group of three small amino acids (Ala Ser Ser), defining a cleavage/binding site for GPI linkage, suggesting that GDNFR- α is an extracellular protein that is attached to the outer cell membrane¹⁵.

To characterize the interaction between GDNF and GDNFR- α , crosslinking and binding experiments were performed. Competition binding of ¹²⁵I-GDNF to Chinese hamster ovary cells stably expressing GDNFR- α demonstrated that GDNF binds specifically and reversibly to GDNFR- α , and that the two proteins interact with an approximate K_d of 63 pM (Fig. 1c). Similarly, when these cells were crosslinked to ¹²⁵I-GDNF, three proteins of apparent relative molecular mass 85,000 (M_r 85K), 180K and 200K were detected (Fig. 1b). These proteins were

absent when the crosslinking reaction took place in the presence of excess unlabelled GDNF, or when ¹²⁵I-GDNF was crosslinked to cells expressing an irrelevant cell-surface protein (Fig. 1b). Crosslinking and equilibrium binding of ¹²⁵I-GDNF, as well as binding of GDNF to GDNFR- α -expressing cells, were virtually abolished after treatment with phosphoinositide-specific phospholipase C (PIPLC), an enzyme that specially cleaves GPI linkage¹⁶ (Fig. 1d and data not shown), supporting the notion that GDNFR- α is indeed a high-affinity GPI-linked, GDNF-binding protein.

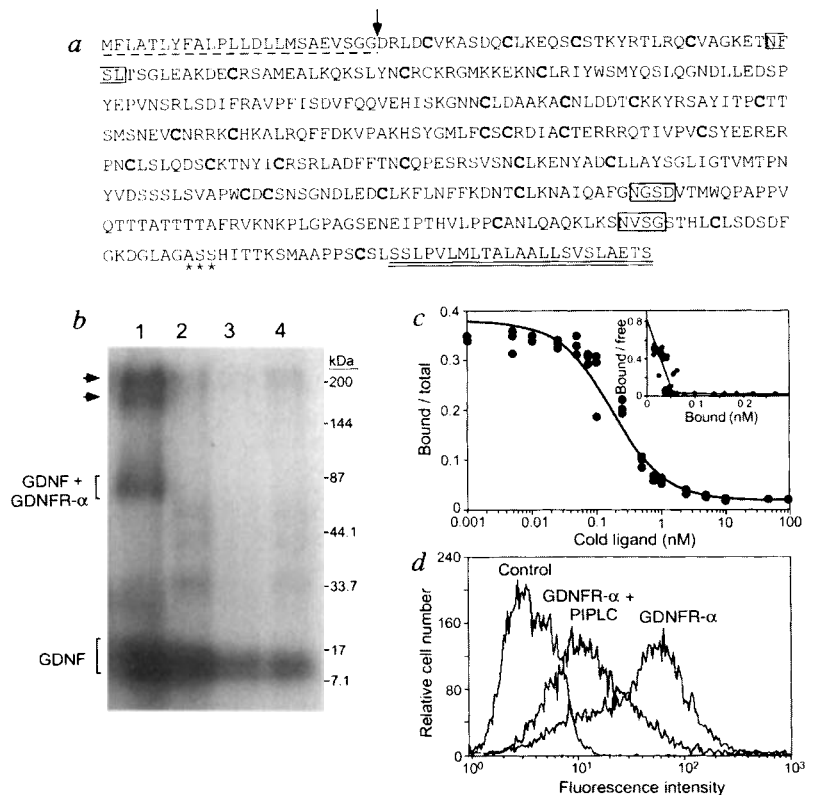
The tissue distribution of the GDNFR- α messenger RNA was examined by northern blot analysis, polymerase chain reaction (PCR) and *in situ* hybridization analysis (Fig. 2 and data not shown). In the embryonic and adult rat nervous system, mRNA for GDNFR- α is found in many GDNF-responsive regions, including ventral midbrain (dopaminergic neurons), ventral spinal cord (spinal motor neurons) and in subpopulations of GDNF-dependent dorsal root ganglia (DRG) neurons. Consistent with the finding that the kidneys and the enteric nervous system fail to develop in GDNF-deficient mice^{7–9}, high levels of GDNFR- α mRNA are present in developing nephrons and in embryonic smooth and striated muscles around the enteric nervous system in the oesophagus, gut and stomach. GDNFR- α transcripts were also found in other regions of the nervous system, including for example the retina, thalamus, pons and medulla oblongata, as well as in non-neuronal tissues including the pituitary, urogenital tract and pancreatic primordium (Fig. 2).

The tissue distribution and biochemical characteristics of GDNFR- α were consistent with its involvement in GDNF signalling. To examine this possibility, populations of primary embryonic neurons known to be dependent on GDNF for survival were treated with PIPLC, and their survival response to GDNF or other factors was monitored in culture. The number of embryonic chick nodose and trigeminal ganglion sensory neurons or sympathetic neurons surviving in the presence of saturating concentrations of GDNF was reduced by 50–70% after PIPLC treatment. No changes in the response of these neurons to brain-derived neurotrophic factor (BDNF) or nerve growth factor (NGF) was observed in the presence of PIPLC (Fig. 3a). Similarly, PIPLC treatment reduced the number of embryonic rat spinal motor neurons or dopaminergic neurons that survived in the presence of GDNF by 50–90% without affecting the survival of these neurons in the presence of BDNF or transforming growth factor (TGF)- β 3 (Fig. 3a). In different systems, PIPLC reduced the survival-promoting effects of GDNF at GDNF concentrations as low as 10 pg ml⁻¹, suggesting that the GPI-linked molecule is necessary for the high-affinity response. PIPLC was effective even when GDNF was applied at 1 μ g ml⁻¹ (2×10^3 -fold above the EC₅₀), excluding the possibility that the GPI-linked protein, following its release from the cell membrane, binds GDNF and reduces its effective concentration (Fig. 3a).

To examine directly the role of GDNFR- α in the response to GDNF, purified motor neurons were treated with PIPLC, and their responses to GDNF and to the combination of GDNF and soluble GDNFR- α were examined. As before, GDNF alone failed to prevent the death of many PIPLC-treated motor neurons. In contrast, when GDNF was added to these motor neurons in combination with soluble GDNFR- α , the response to GDNF was restored (Fig. 3b). To confirm further that GDNFR- α is an important mediator of GDNF signals, rat pheochromocytoma PC12 cells, which are dependent on neurotrophic factors for survival in serum-free media and express low levels of Ret (data not shown), were grown without serum in the presence of GDNF, soluble GDNFR- α or both, and examined 7 days later. With GDNF or soluble GDNFR- α alone, only a few neurite-bearing, phase-bright, live cells were found. In contrast, when PC12 cells were cultured with both GDNF and GDNFR- α , an increase in the number of living cells with neurites was observed (Fig. 3c). GDNFR- α is thus an important component of the GDNF signalling cascade, and has the properties expected of the ligand-binding subunit of a functional GDNF receptor.

FIG. 1 Primary structure and GDNF-binding characteristics of GDNFR- α . **a**, Protein sequence of the GDNFR- α . The signal peptide is indicated by a broken line, the putative signal cleavage site is marked with an arrow, potential glycosylation sites are boxed, the hydrophobic domain of the GPI attachment site is doubly underlined, the small amino-acid residues that constitute the cleavage/attachment site for GPI-linked proteins are marked with asterisks, and the cysteine residues are indicated in bold type. In addition to the cDNA isolated by expression cloning, 9 other cDNAs were isolated from rat (4) and mouse (5) cDNA libraries using GDNFR- α cDNA as a probe; 8 contained an open reading frame (ORF) identical to GDNFR- α , but one rat cDNA encoded a shorter ORF of 158 amino acids, which may represent an aberrant or a secreted form of this protein. **b**, Crosslinking of 125 I-GDNF to cells expressing GDNFR- α (lanes 1, 2) or to control cells (lanes 3, 4) in the absence (lanes 1, 3) or presence (lanes 2, 4) of excess unlabelled GDNF. Radioactive bands (~ 85 K, ~ 180 K, ~ 200 K), which can be displaced by unlabelled GDNF, are present in GDNFR- α -expressing, but not in PIPLC-treated (not shown) or control, cells. The protein band at 80–85K is likely to represent a complex of 58K GDNFR- α and the 15K GDNF monomer, whereas the bands at higher M_r may represent interaction between the 125 I-GDNF, GDNFR- α and putative signalling molecules such as Ret (see below), or dimerization of the 125 I-GDNF/GDNFR- α complex. **c**, Binding of 125 I-GDNF to GDNFR- α -expressing cells and displacement by unlabelled GDNF. The Scatchard analysis (inset) gave a K_d value of 63 pM. **d**, Flow cytometry analysis of the effects of PIPLC on GDNF binding to cells expressing GDNFR- α . Control, cells expressing an irrelevant cell surface protein; GDNFR- α +, PIPLC cells expressing GDNFR- α treated for 1 h at 37 °C with 2 $\mu\text{g ml}^{-1}$ PIPLC; GDNFR- α -, cells expressing GDNFR- α that were not treated with PIPLC. A shift to the right indicates binding to GDNF. Treatment of GDNFR- α -expressing cells with PIPLC leads to a reduction of over 90% in the amount of GDNF binding.

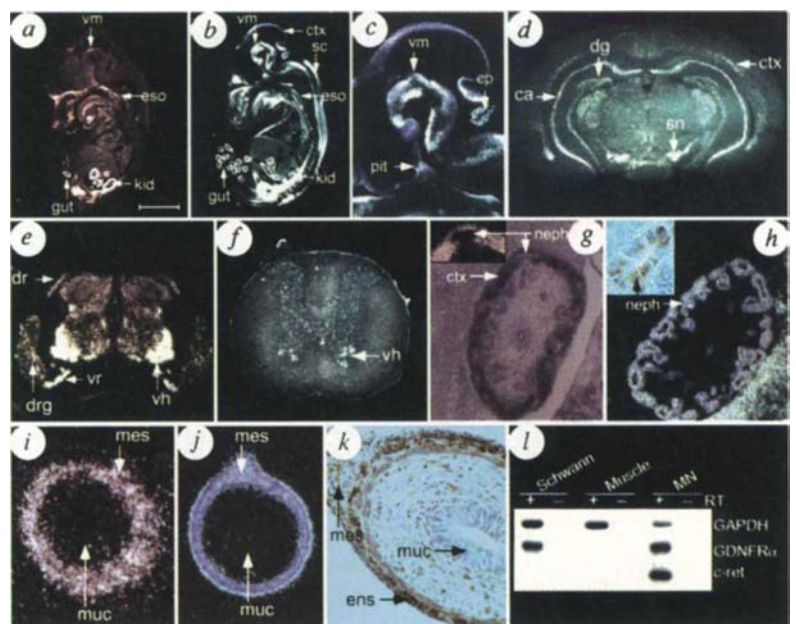
METHODS. For crosslinking, Chinese hamster ovary (CHO) cells stably expressing GDNFR- α or an irrelevant protein were incubated for 1 h at 37 °C either in the presence or absence of PIPLC (2 $\mu\text{g ml}^{-1}$) and were then resuspended at a density of 10^6 to 2×10^6 per ml in ice-cold L15 media with 1 mM phenylmethylsulphonyl fluoride, 1 mg ml^{-1} B.S.A. and 50 pM 125 I-labelled GDNF, and incubated at 4 °C for 2 h. Formaldehyde was added to a final concentration of 4% at room temperature for 30 min, and the cells were washed 3 times with 1 ml phosphate-buffered saline. Cells were then lysed in sample buffer (80 mM Tris-HCl, pH 6.8, 10% (v/v) glycerol, 1% (w/v) SDS, 0.025% bromophenol blue and loaded on to SDS-polyacrylamide gels. For equilibrium binding analysis, cells were processed as before and incubated with



10 pM 125 I-labelled GDNF and various concentrations of cold GDNF. The IGOR program was used to determine K_d . For flow cytometry analysis, cells were incubated for 1 h at 37 °C either in the presence or absence of PIPLC (2 $\mu\text{g ml}^{-1}$). GDNF (100 ng ml^{-1}) and anti-GDNF monoclonal antibodies (100 $\mu\text{g ml}^{-1}$) were then applied. The cells were washed in ice-cold PBS and were incubated for an additional 30 min with fluorescent anti-IgG monoclonal antibodies (Vector), washed, and the cells analysed using a flow cytometer. PIPLC was produced and purified as described¹⁶.

FIG. 2 Tissue distribution of GDNFR- α compared to GDNF and Ret. **a**, **b**, Sagittal sections through embryonic day 15 (E15) rat embryo. **c**, Sagittal section through an E15 rat midbrain. **d**, Coronal section of an adult rat brain at the level of substantia nigra. **e**, Transverse section of E15 rat spinal cord. **f**, Transverse section through an adult rat spinal cord. **g**, **h**, E15 rat kidney. **i**, **j**, E15.5 rat gut. **a**, **g**, **i**, *In situ* hybridization with a GDNF probe. **b**–**f**, **h**, **j**, *In situ* hybridization with a GDNFR- α probe. **k**, E15.5 rat gut stained with antisera to Ret. **l**, Reverse transcription (RT)-PCR analysis of: Schwann, Schwann cell line MSC80; Muscle, adult soleus muscle; MN, purified motor neurons from E14 rat. Motor neurons express both GDNFR- α and c-ret. The inset in **h** represents immunohistochemical staining of a developing nephron with Ret antiserum. In the kidney both Ret and GDNFR- α are expressed in the developing nephrons adjacent to GDNF. In the gut, GDNF and GDNFR- α are present between the inner circular and outer longitudinal smooth muscle adjacent to and possibly within the enteric nervous system (ENS), whereas Ret is present only in the ENS. In addition to the tissues listed in the text, GDNFR- α transcripts are found in the E15.5 rat outer layer of the midbrain tectum, choroid plexus, cerebellar primordium, the olfactory epithelium, whisker pads, genital tubercle, urogenital sinus, testes, the intervertebral discs and trachea. In the adult, GDNFR- α transcripts were found in the pars compacta region of the substantia nigra, the ventrolateral cell column of the spinal cord, the hippocampal formation, inner layers of the cerebral cortex, lateral geniculate nucleus, superior colliculus, outer margin of cerebellar granule cells, lateral septum, endopiriform nucleus, and claustrum. Abbreviations: vm, ventral midbrain; eso, oesophagus; kid, kidney; ctx, cortex; pit, pituitary; sc, spinal cord; cp, choroid plexus; sn, substantia nigra; dg, dentate gyrus; ca, cornu ammonis; dr, dorsal root; drg, dorsal root ganglia; vr, ventral root; vh, ventral horn; neph, nephron; mes, mesentery; muc, mucosal epithelia. Scale bar: **a**, **b**, 2.5 mm; **c**, 3 mm; **d**, 2.1 mm; **e**, 250 μm ; **f**, 700 μm ; **g**, **h**, 200 μm ; **i**, **j**, 100 μm ; **k**, 6 μm .

METHODS. For *In situ* hybridization, E15.5 rat embryos were immersion-fixed overnight at 4 °C in 4% paraformaldehyde, then cryoprotected overnight in 15% sucrose.



Adult rat brains and spinal cords were frozen fresh. All tissues were sectioned at 16 μm , and processed for *in situ* hybridization using ^{33}P -UTP labelled RNA probes as described⁵. Sense and antisense probes were derived from the N-terminal region of GDNFR- α using T7 polymerase. RT-PCR analysis was performed as described⁵.

Because GDNFR- α is anchored to the outer surface of the cell, transmission of GDNF signals following binding to GDNFR- α must involve an additional transmembrane protein. The striking similarity between the phenotype of the GDNF-deficient mice⁷⁻⁹ and mice lacking the orphan tyrosine kinase receptor *c-Ret*^{17,18}, as well as the tissue distribution of Ret¹⁹⁻²¹ (Fig. 2*h, i, k*), raised the possibility that GDNF and Ret might function in the same signalling cascade. To examine this hypothesis, the human neuro-

blastoma SK-N-SH and the mouse neuroblastoma Neuro-2a, cell lines that express endogenous *c-ret*, were exposed to GDNF alone or to GDNF in combination with soluble GDNFR- α for 5 min, and the level of Ret tyrosine phosphorylation was determined. GDNF induced modest phosphorylation of Ret in these two cell lines (Fig. 4*a*) but not in NIH3T3 cells that stably express the human Ret (data not shown). Ret phosphorylation was further increased when GDNF was added together with GDNFR- α , but

FIG. 3 GDNFR- α is required for survival responses to GDNF. *a*, Survival response of embryonic chick nodose, trigeminal, sensory and sympathetic neurons, rat spinal motor neurons and rat dopaminergic neurons to GDNF or to other growth factors following treatment with PIPLC. PIPLC reduces cell survival in the presence of GDNF or CNTF by 50–90% without changing the response to BDNF, NGF or TGF- β 3. *b*, Soluble GDNFR- α (sR α) restores the response of PIPLC-treated motor neurons to GDNF. The trophic activity of GDNFR- α alone these experiments is believed to be due to the low levels of GDNF thought to be associated with this preparation. *c*, Soluble GDNFR- α (sR α) is required to elicit a response of PC12 cells to GDNF. The number of neurites bearing live cells per microscopic field is presented.

METHODS. Embryonic chick nodose, trigeminal and sympathetic ganglia neurons², E14 rat motoneurons⁵, and E14 rat dopaminergic neurons²⁸ were isolated, plated and grown in triplicate wells as described. PIPLC (2–4 μ M ml⁻¹) was added to the indicated samples 1–2 h before, and 12 h and 24 h after, addition of the indicated growth factors, and the number of surviving neurons was determined 30 and 72 h later. PC12 cells were seeded on collagen polyornithine-coated 35-mm dishes in RPMI medium supplemented with 10% horse serum and 5% fetal calf serum. After attachment, the cells were switched to serum-free medium and then exposed to GDNF (50 ng ml⁻¹) and soluble GDNFR- α (sR α) as indicated. The number of live neurite-bearing phase-bright cells per microscopic field was determined 7 days later as described²⁹. Soluble GDNFR- α was produced as a C-terminus His-tagged protein in 293 human embryonic kidney cells, was purified by Ni-NTA chromatography as described³⁰, and was tested for GDNF binding.

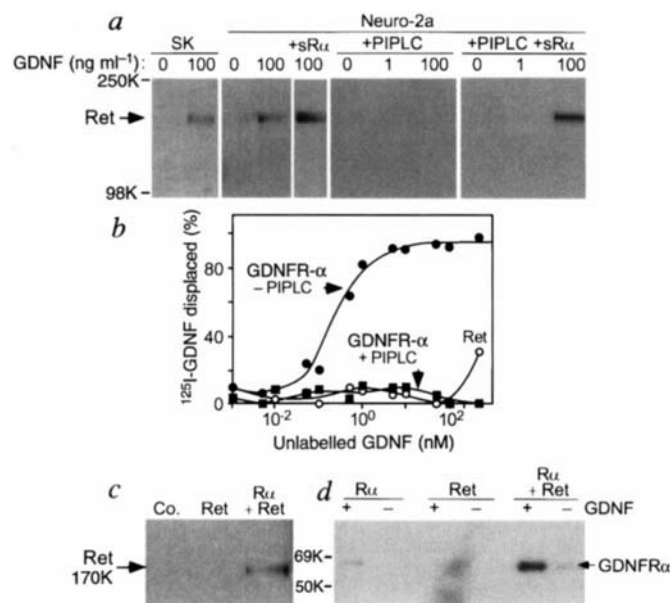
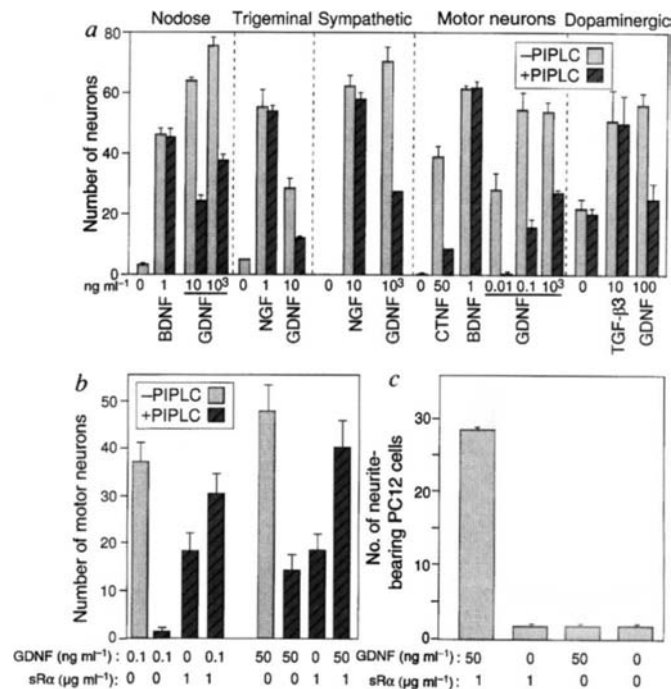


FIG. 4 Involvement of Ret in the response to GDNF. *a*, GDNF in conjunction with GDNFR- α (sR α) induces tyrosine phosphorylation of Ret. Modest stimulation of Ret tyrosine phosphorylation was observed in Neuro-2a and SK-N-SH (SK) cells that were not treated with PIPLC following exposure to GDNF alone (left 2 panels). Phosphorylation of Ret was further enhanced in the presence of soluble GDNFR- α (+sR α). No stimulation of Ret phosphorylation was observed in PIPLC-treated cells (PIPLC) unless GDNF was added together with GDNFR- α (PIPLC + sR α). *b*, Binding of ¹²⁵I-GDNF to cells expressing GDNFR- α or Ret. GDNF does not bind Ret with high affinity. *c*, GDNF, GDNFR- α and Ret from a complex on the cell surface. Co,

untransfected cells; Ret, cells transfected with Ret alone; R α + Ret, cells transfected with Ret and GDNFR- α . In all cases cells were exposed to GDNF (100 ng ml⁻¹) and then processed for immunoprecipitation with GDNF antisera. The presence of immune complexes between GDNF and Ret was then determined on a western blot with Ret antisera. GDNF/Ret complex was formed only in the presence of GDNFR- α . *d*, Formation of GDNFR- α /Ret is stimulated by GDNF. R α , cells transfected with an epitope-tagged GDNFR- α alone. Ret, cells transfected with Ret alone R α + Ret, cells transfected with Ret and an epitope-tagged GDNFR- α . After transfection, cells were either treated with GDNF (+) or left untreated (–) and then processed for immunoprecipitation with Ret antisera. The presence of immune complexes between Ret and GDNFR- α was then determined on a western blot with antisera to the epitope-tagged GDNFR- α . Immune complexes between GDNFR- α and Ret were formed in the presence of GDNF.

METHODS. To assay for tyrosine phosphorylation, cells were incubated for 1 h at 37 °C with or without PIPLC, and then exposed to various concentrations of GDNF and soluble GDNFR- α for 5–10 min at 37 °C. Cells were then removed from the plates with 2 mM EDTA in PBS and lysed with ice-cold buffer (10 mM sodium phosphate, pH 7.0, 100 mM NaCl, 1% NP40, 5 mM EDTA, 100 mM sodium vanadate, 2 mM PMSF, and 0.2 units of aprotinin) and used for immunoprecipitation with antisera raised against the 19-amino-acid C terminus of Ret followed by binding to protein A sepharose. The immunoprecipitated proteins were released by boiling in SDS sample buffer, separated on an 8% SDS-polyacrylamide gel, transferred to a nitrocellulose membrane and reacted with anti-phosphotyrosine antibody (Upstate Biotechnology); detection was with an ECL western blotting detection system (Amersham). To increase the level of Ret, SK-N-SH cells were treated with 10 nM retinoic acid 12 h before addition of GDNF. To examine the formation of protein complexes, 293 human kidney cells were transiently transfected with expression vectors for GDNFR- α , an epitope-tagged GDNFR- α , Ret or a combination of expression vectors. Cells were stimulated with GDNF as indicated and lysed with brij 96 detergent (Sigma) as described²². Putative immune complexes were immunoprecipitated with a polyclonal antibody against GDNF (*c*) or Ret (*d*), transferred onto a nitrocellulose filter, and analysed with polyclonal antibody against Ret (*c*) or the epitope-tagged GDNFR- α (*d*).

not when GDNFR- α was added alone (Fig. 4a and data not shown). To determine whether induction of Ret tyrosine phosphorylation was dependent on the presence of GDNFR- α , Neuro-2a and SK-N-SH cells were treated with PIPLC, and the response of Ret to GDNF was examined. In agreement with the finding that survival responses to GDNF require the presence of GDNFR- α , no induction of tyrosine phosphorylation on Ret was detected in these PIPLC-treated cells in the presence of GDNF alone. In contrast, stimulation of tyrosine phosphorylation of the 170K Ret protein was readily observed in PIPLC-treated Neuro-2a and SK-N-SH cells when GDNF was added together with a soluble GDNFR- α (Fig. 4a and data not shown).

Although GDNF stimulated tyrosine phosphorylation of Ret, no high-affinity binding of GDNF to Ret could be detected in Neuro-2a cells expressing high levels of an endogenous Ret, or in cells expressing recombinant Ret protein (Fig. 4b and data not shown). We therefore determined whether physical interactions between Ret and GDNF, as defined by the formation of an immunoprecipitable Ret/GDNF complex, could be mediated by GDNFR- α . Human embryonic kidney 293 cells were transiently transfected with an expression vector containing *c-ret*, or with a combination of expression vectors for *c-ret* and GDNFR- α , treated with GDNF, and then lysed with a mild detergent²². Proteins that formed complexes with GDNF were immunoprecipitated with a polyclonal antibody to GDNF and analysed on a western blot using a polyclonal antibody to Ret. In cells expressing Ret alone or GDNFR- α alone, no co-immunoprecipitated Ret protein could be detected. In contrast, Ret was readily co-immunoprecipitated by GDNF antibodies from cells that express both Ret and GDNFR- α (Fig. 4c). To characterize further the complex between GDNFR- α and Ret, 293 cells were transfected with expression vectors for *c-ret* and with an epitope-tagged GDNFR- α , and then analysed for the presence of GDNFR- α /Ret protein complexes in the presence or absence of GDNF. In cells that expressed Ret alone or the epitope-tagged GDNFR- α alone, no significant level of protein complexes could be detected, either in the presence or absence of GDNF (Fig. 4d). In contrast, in cells expressing both epitope-tagged GDNFR- α and Ret, a protein complex containing both proteins was readily detected after exposure to GDNF (Fig. 4d). These findings are consistent with the ideas: that GDNF, GDNFR- α and Ret can, in the presence of GDNF, form a complex on the cell surface; that Ret is a component of a functional GDNF receptor; and that GDNFR- α is an essential intermediary in the interaction between GDNF and Ret.

In summary, we have provided evidence that GDNF uses a unique receptor system in which GDNFR- α , a novel GPI-linked protein, is a ligand-binding component and the tyrosine kinase receptor Ret is a signalling component. We cannot exclude the possibility that the GDNF receptor complex contains components other than GDNFR- α and Ret, or that other types of GDNF receptors exist. The GDNF receptor complex resembles in part the receptors for ciliary neurotrophic factor²², bacterial endotoxin^{23,24}, and receptors in the immune system in which GPI-linked proteins serve as ligand-binding components and cytoplasmic tyrosine kinases serve as the signalling components²⁵. Similarly, a receptor system involving the novel tyrosine kinase receptor Musk and an as yet unidentified accessory molecule, was shown to transduce Agrin signals in the neuromuscular junction²⁶. GDNF might represent an evolutionary transition within the superfamily of the cysteine-knot-containing proteins, which includes growth factors that use serine threonine kinase receptors (the TGF- β branch of this superfamily), and growth factors that use tyrosine kinase receptors (the nerve growth factor and platelet-derived growth factor branches of this superfamily)²⁷. □

- Beck, K. D. et al. *Nature* **373**, 339–341 (1995).
- Tomac, A. et al. *Nature* **373**, 335–359 (1995).
- Henderson, C. E. et al. *Science* **266**, 1062–1064 (1994).
- Arenas, E., Trupp, M., Akerud, P. & Ibáñez, C. F. *Neuron* **15**, 1465–1473 (1995).
- Pichel, J. S. et al. *Nature* **382**, 73–76 (1996).
- Sanchez, M. P. et al. *Nature* **382**, 70–73 (1996).
- Moore, M. W. et al. *Nature* **382**, 76–79 (1996).
- Takahashi, M., Ritz, J. & Cooper, G. M. *Cell* **42**, 581–588 (1985).
- Takahashi, M. & Cooper, G. M. *Molec. cell. Biol.* **7**, 1378–1385 (1987).
- Takahashi, M. *Crit. Rev. Oncogen.* **6**, 35–46 (1995).
- Gearing, D. P., King, J. A., Gough, N. M. & Nicola, N. A. *EMBO J.* **8**, 3667–3676 (1989).
- Bazan, J. F. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6934–6938 (1990).
- Micanovic, R., Gerber, L. D., Berger, J., Kodukula, K. & Udenfriend, S. *Proc. natn. Acad. Sci. U.S.A.* **87**, 157–161 (1990).
- Koke, J. A., Yang, M., Henner, D. J., Volwerk, J. J. & Griffith, O. H. *Prot. Express. Purif.* **2**, 51–58 (1991).
- Schuchardt, A., D'Agati, V., Larsson-Blomberg, L., Costantini, F. & Pachnis, V. *Nature* **367**, 380–383 (1994).
- Durbec, P. L., Larsson-Blomberg, L. B., Schuchardt, A., Costantini, F. & Pachnis, V. *Development* **122**, 349–358 (1996).
- Pachnis, V., Mankoo, B. & Costantini, F. *Development* **119**, 1005–1007 (1993).
- Avantaggiato, V. et al. *Cell Growth Diff.* **5**, 305–311 (1994).
- Tsuzuki, T. et al. *Oncogene* **10**, 191–198 (1995).
- Davis, S. et al. *Science* **259**, 1736–1739 (1993).
- Lee, J.-D. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 9930–9934 (1993).
- Pugin, J. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 2744–2748 (1993).
- Brown, D. *Curr. Opin. Immunol.* **5**, 349–354 (1993).
- Glass, D. J. et al. *Cell* **85**, 513–523 (1996).
- McDonald, N. Q. & Hendrickson, W. A. *Cell* **73**, 421–424 (1993).
- Poulsen, K. T. et al. *Neuron* **13**, 1245–1252 (1994).
- Loeb, D. M. et al. *Cell* **66**, 961–966 (1991).
- Attisano, L., Wrana, J. L., López-Casillas, F. & Massagué, J. *Biochim. biophys. Acta* **1222**, 21–80 (1994).

ACKNOWLEDGEMENTS. J.J.S.T., L.G. and F.d.S. contributed equally to this work. We thank W. Anstine for preparing the figures; B. Zalc for the MSC80 line; E. Berry for help with the manuscript; G. Bennett for help with the protein iodination and binding assays; A. Ryan for help with the immunohistochemistry; H. Parshad and R. Kiss for help in protein production; M. Vasser, P. Jhurani and P. Ng for synthetic oligonucleotides; M. Yang for the PIPLC enzyme; V. Arce for help with the PIPLC experiments; and M. Hynes for critical reading of the manuscript. This work was supported in part by INSERM, Association Française contre les Myopathies, and Institut pour la Recherche sur la Moelle Epinière, and by a fellowship from Wellcome Trust (R.A.P.) and by a grant from Action Research (A.M.D.).

CORRESPONDENCE and requests for materials should be addressed to A.R. (e-mail: ar@gene.com).

Regulation by cAMP-dependent protein kinase of a G-protein-mediated phospholipase C

Mingyao Liu & Melvin I. Simon

Division of Biology, 147-75, California Institute of Technology, Pasadena, California 91125, USA

THE heterotrimeric G proteins mediate a variety of cellular processes by coupling transmembrane receptors to different effector molecules, including adenylyl cyclases and inositol-phospholipid-specific phospholipase C (PLC)^{1–3}. Activation of adenylyl cyclases results in the production of cyclic AMP and activation of cAMP-dependent protein kinase (PKA). Phospholipase C catalyses the hydrolysis of phosphatidylinositol-4,5-bisphosphate (PtdInsP₂) to generate diacylglycerol and inositol-1,4,5-trisphosphate (InsP₃), leading to the activation of protein kinase C (PKC) and the mobilization of intracellular calcium⁴. The various PLC isoforms appear to be activated by different receptors, and in some cases by different G-protein components⁵. There are four well-characterized forms of PLC- β and all of them are activated to various extents by the G α_q family of G proteins^{6–9}. Specific activation of PLC isoforms β_2 and β_3 by G-protein $\beta\gamma$ subunits has also been reported^{10,11}. Although it has been suggested that PLC activity might be modulated by the adenylyl cyclase pathway, no clear link has been established between the two pathways. Here we report that cAMP-dependent protein kinase specifically inhibits G $\beta\gamma$ -activated PLC- β_2 activity but not that of the G α -activated PLC isoforms, and that the effect of PKA is not mimicked by PKC isozymes. Furthermore, we show that PKA directly phosphorylates serine residues of the PLC- β_2 protein both *in vivo* and *in vitro*. Our results provide an insight into the

Received 24 May; accepted 14 June 1996.

- Lin, L. H., Doherty, D. H., Lile, J. D., Bektesh, S. & Collins, F. *Science* **260**, 1130–1132 (1993).
- Buj-Bell, A., Buchman, V. L., Horton, A., Rosenthal, A. & Davies, A. M. *Neuron* **15**, 821–828 (1995).

specificity and nature of the crosstalk between the two G-protein-coupled signal transduction pathways.

To examine whether cAMP-dependent protein kinase is involved in regulating phospholipase C activity, we reconstituted the G-protein $\beta\gamma$ -stimulated PLC signalling pathway in COS-7 cells in the presence or absence of the catalytic subunit of PKA¹². Specific activation of PLC- $\beta 2$ by G $\beta\gamma$ subunits was demonstrated by measuring the release of inositol phosphates resulting from the hydrolysis of PtdInsP₂. Co-transfection of complementary DNAs encoding PLC- $\beta 2$ and G $\beta_5\gamma_2$ subunits leads to ~5-fold stimulation of PLC activity (Fig. 1a). However, upon cotransfection with a cDNA encoding the catalytic subunit of PKA, G $\beta\gamma$ -stimulated PLC activity was totally inhibited (Fig. 1a). There was little or no effect after transfection of PLC- $\beta 2$ and G-protein $\beta\gamma$ subunits alone. The possible targets for PKA inhibition of the PtdInsP₂ hydrolysis include the G-protein $\beta\gamma$ subunits, the PLC- $\beta 2$ protein, or both. Using other $\beta\gamma$ complexes ($\beta_1\gamma_2$ and $\beta_2\gamma_2$) of G proteins that are not phosphorylated *in vitro* by PKA (our unpublished results), we observed similar inhibitory effects of PKA on PLC- $\beta 2$

activity (Fig. 1b), suggesting that PKA phosphorylates PLC- $\beta 2$ rather than the G $\beta\gamma$ subunits. To understand better the nature of protein kinase action, we also examined the effects of three different protein kinase C isozymes¹³, α , β (both are Ca²⁺ and diacylglycerol-dependent) and ϵ (only diacylglycerol-dependent), on G $\beta\gamma$ -activated PLC- $\beta 2$ activity. As shown in Fig. 1c, stimulated PKC isozymes have little effect on PLC- $\beta 2$ activity, correlating well with a previous report that phosphorylation of PLC- β by PKC has no concomitant effect on the activity of PLC¹⁴. These results demonstrate that the regulatory effect of PKA on PLC- $\beta 2$ is specific and not mimicked by other protein kinases.

The inhibitory effect of PKA on other PLC- β isoforms was investigated by measuring receptor-coupled G α -activated endogenous PLC- β activity in COS-7 cells. cDNAs encoding different receptors, G α_q isoforms (G α_{15} or G α_{16}), and the PKA catalytic subunit were cotransfected into the cells as indicated. G α_{15} and G α_{16} couple a variety of receptors to the endogenous PLC- β isoforms in COS-7 cells¹⁵. Figure 2a shows the coexpression of the β_2 -adrenergic receptor with G α_{15} or G α_{16} leads to a 4–5-fold

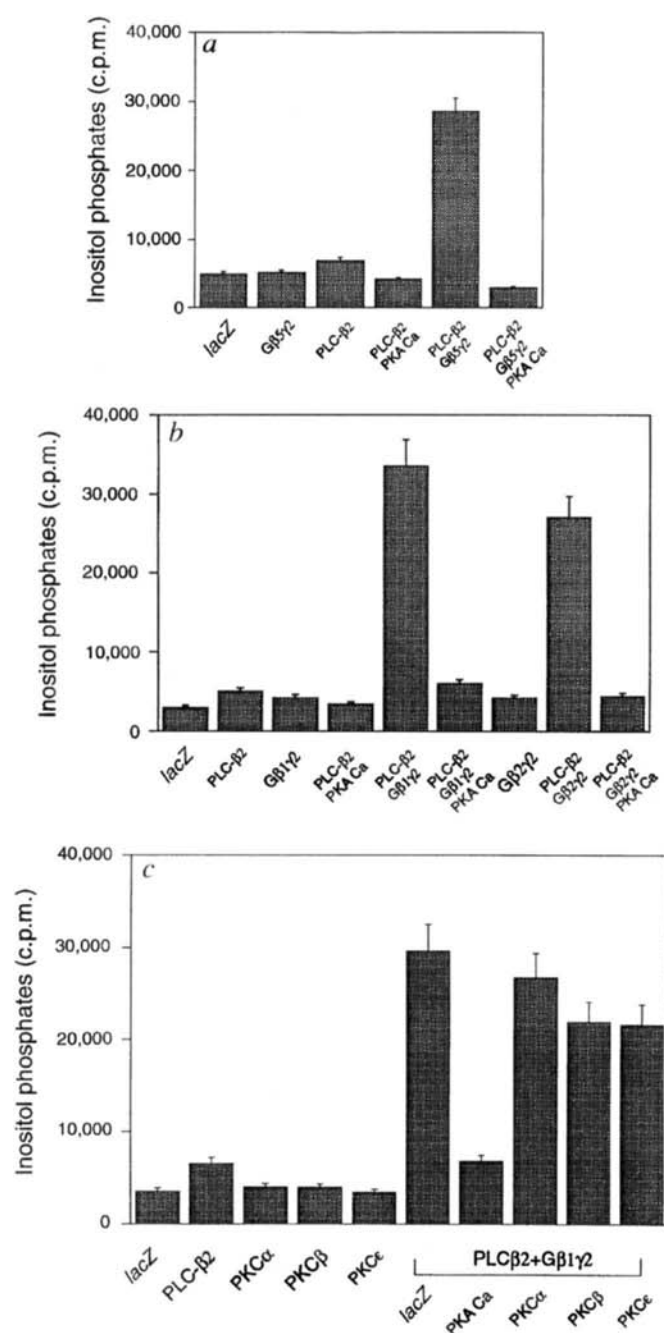


FIG. 1 Regulation of G $\beta\gamma$ -activated PLC- $\beta 2$ by protein kinases in COS-7 cells. **a** and **b**, Inhibitory effect of PKA on G $\beta\gamma$ -stimulated PLC- $\beta 2$ activity. Accumulation of inositol phosphates was measured in COS-7 cells transfected with expression vectors carrying cDNAs corresponding to β -galactosidase (*lacZ*), PLC- $\beta 2$, G β_1 , G β_2 , G β_5 , G γ_2 and PKA catalytic subunit (Ca). The cDNAs transfected are indicated beneath each column. **c**, Effect of protein kinase C isozymes (α , β and ϵ) on G $\beta\gamma$ -stimulated PLC- $\beta 2$ activity. Transfection of PKC isoforms into COS-7 cells increases unstimulated PKC activity ~1.5-fold, but PKC activity further ~3-fold upon stimulation with TPA (12-O-tetradecanoylphorbol-13-acetate). Peptides IYRRGSRRLRLK and MNRRGSIKQAKI were used as substrates for PKC isoforms. Permeabilized COS-7 cells were used for the assay. Data represent duplicate determinations in a single experiment; error bar gives the range of duplicates. Results were similar in three additional experiments. The relative levels of each expressed component in transfections were similar, as determined by western blotting (data not shown). **METHODS.** COS-7 cells (1×10^5 cells per well) were seeded in 12-well plates in Dulbecco's modified Eagle medium (DMEM) containing 10% fetal bovine serum (FBS) 24 h before transfection. The total amount of DNA in all transfections was 1 μ g per well. A cytomegalovirus vector pCIS encoding β -galactosidase was used to maintain a constant amount of DNA and equalize the amount of a particular cDNA in each set of experiments. Cells were transfected as described^{11,15}. Briefly, cDNAs were mixed with lipofectamine (BRL) in serum-free Opti-MEM (BRL) and cells were transfected for 5 h; 24 h later, cells were washed with PBS and incubated in 0.5 ml inositol-free DMEM containing 10 μ Ci ml⁻¹ of [³H]myo-inositol (Amersham) and 10% dialysed FBS. 24 h later, cells were washed with PBS, incubated in inositol-free medium containing 10 mM LiCl for 25 min at 37 °C, then lysed in ice-cold 10 mM formic acid for 30 min; the lysate was neutralized with 15 mM NH₄OH. Total inositol phosphates were separated on an anion-exchange column (AG 1-X8, Bio-Rad)²⁶; the eluted inositol phosphates were mixed with 10 ml scintillation cocktail (BCS; Amersham) and counted.

increase of endogenous PLC activity upon agonist stimulation. But unlike $G\beta\gamma$ -activated PLC- β_2 , cotransfection of the PKA catalytic subunit has no significant effect on the ligand-dependent, $G\alpha_q$ -stimulated endogenous PLC- β activity in COS-7 cells. These cells primarily express the β_1 and β_3 isoforms of PLC¹¹. Similar results were also obtained when other receptors and the $G\alpha_q$ isoforms were used in cotransfection assays (for example, the M1 muscarinic receptor; data not shown). Therefore, PKA seems to inhibit only $G\beta\gamma$ -activated PLC- β_2 , and not the activation of $G\alpha$ -coupled PLC isoforms in COS-7 cells (Fig. 2a, b). To investigate further the functional effect of PLC- β_2 phosphorylation, we also examined the effect of forskolin, a stimulator of adenylyl cyclase, on agonist-stimulated PtdInsP₂ hydrolysis in human HL60 promyelocytic leukaemia cells. Differentiated HL60 cells express both endogenous fMet-Leu-Phe (fMLP) receptors and the PLC- β_2 enzyme. Upon agonist stimulation, activated fMLP receptors interact with $G\alpha_i$ subfamily proteins, and the released $G\beta\gamma$ subunits activate PLC- β_2 in a pertussis-toxin (PTX)-sensitive manner¹⁶. When differentiated HL60 cells were preincubated with forskolin, the fMLP-stimulated PLC activity was largely inhibited, as shown by the decrease in inositol phosphate production (Fig. 2c); this suggests that the increase in cAMP caused by forskolin probably activates PKA and so inhibits $G\beta\gamma$ -stimulated PLC- β_2 . This is in agreement with earlier reports about the inhibitory effects of cAMP on agonist-stimulated hydrolysis of PtdInsP₂ in several specific cells and tissues^{17–21}. Together, these data support our observations that cAMP-dependent protein kinases regulate $G\beta\gamma$ -stimulated PLC- β_2 activity in the transfection assays. The $G\beta\gamma$ activation of PLC- β_2 may be the primary pathway that accounts for the PTX-sensitive activation of PLC in some cells and tissues.

Cotransfection of a cDNA encoding the PKA catalytic subunit into COS-7 cells has little effect on the expression level of PLC- β_2 . Expression of this protein was confirmed by immunoblotting using anti-peptide antibodies directed against the carboxy-terminal 15 amino acids of PLC- β_2 . These antibodies detected a major band of relative molecular mass ~140K of comparable intensity (Fig. 3a) in the presence and absence of the PKA catalytic subunit, indicating that the effect of PKA on PLC activity is not due to differences in PLC- β_2 expression.

We investigated whether PLC- β_2 was directly phosphorylated by PKA using *in vivo* labelling and *in vitro* phosphorylation experiments. PLC- β_2 was expressed in COS-7 cells, prelabelled with either [³⁵S]methionine or [³²P]orthophosphate, and isolated by immunoprecipitation with anti-PLC- β_2 antibodies. A 140K protein was precipitated by the specific antibody and meta-

bolically labelled with both ³⁵S and ³²P in cells transfected with a cDNA encoding PLC- β_2 (Fig. 3b). This protein was not detected in COS-7 cells transfected with a control plasmid (*lacZ*), nor was it precipitated by a control antiserum. The labelling of PLC- β_2 in the absence of PKA catalytic subunit reveals that this protein has a low basal phosphorylation (Fig. 3b, lane 1), perhaps due to endogenous protein kinase. However, coexpression of PLC- β_2 with the catalytic subunit of PKA resulted in a large increase in the phosphorylation of PLC- β_2 (Fig. 3b, lane 2), indicating that PKA probably directly phosphorylates this protein. Treatment of HL60 cells that express endogenous PLC- β_2 with forskolin (100 μ M) also increased the phosphorylation of PLC- β_2 (data not shown). To confirm that PLC- β_2 is phosphorylated by PKA, we isolated and purified PLC- β_2 by immunoprecipitation and then phosphorylated the protein *in vitro* using purified PKA catalytic subunits in the presence of [³²P]ATP. Figure 3c shows the time-dependent phosphorylation of PLC- β_2 , which reached a plateau after 5 min at 37°C. These results are evidence that PKA rapidly phosphorylates PLC- β_2 .

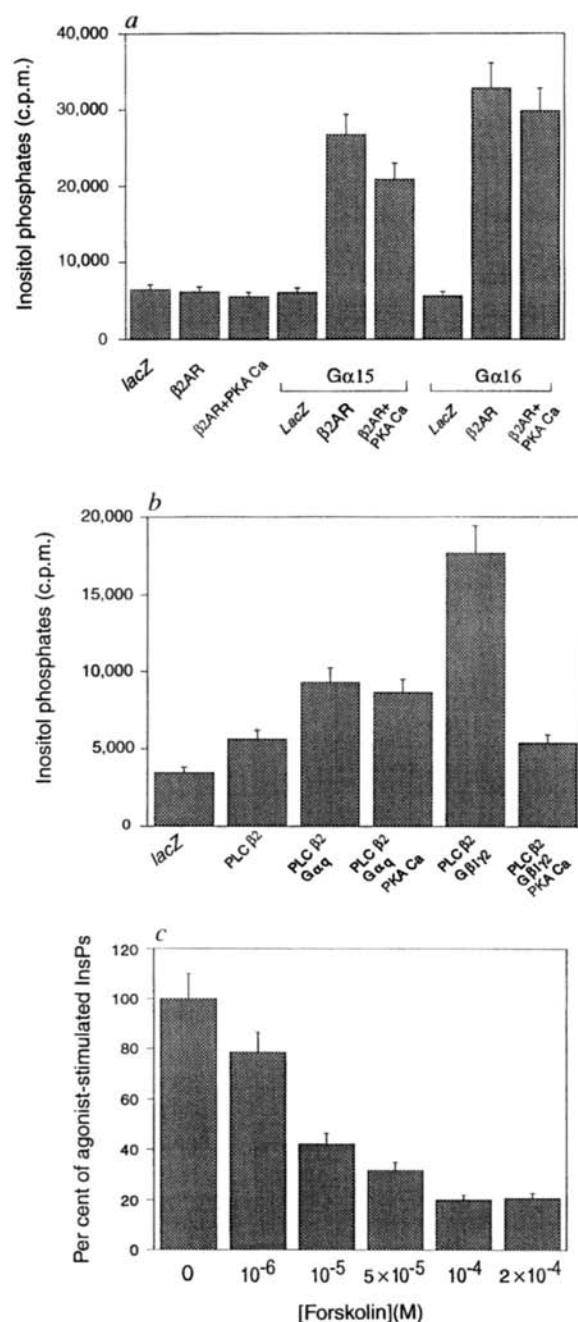


FIG. 2 a, Effect of PKA on the accumulation of ligand-dependent inositol phosphates in COS-7 cells coexpressing β_2 -adrenergic receptor and $G\alpha_{15}$ or $G\alpha_{16}$ subunits. COS-7 cells were cotransfected with cDNAs encoding the β_2 -adrenergic receptor (β_2AR), β -galactosidase (*lacZ*), and subunits of the $G\alpha_q$ family ($G\alpha_{15}$ and $G\alpha_{16}$), in the cytomegalovirus vector pCIS. After 48 h transfection, [³H]inositol-labelled cells were incubated for 20 min with 20 μ M isoprenaline. Inositol phosphates were determined as described¹⁵. b, Lack of effect of PKA Ca on $G\alpha_q$ -activated PLC- β_2 . Coexpression of $G\alpha_q$ with PLC- β_2 resulted in ~1.5-fold increase in PLC activity. Whereas PKA dramatically reduced $G\beta\gamma$ -stimulated PLC- β_2 activity, PKA had little effect on the $G\alpha_q$ -stimulated PLC- β_2 activity. Details and methods as for Fig. 1. c, Inhibition of fMLP-stimulated inositol phosphates in differentiated HL60 cells by the adenylyl cyclase activator forskolin. HL60 cells were grown in RPMI 1640 medium supplemented with 10% FBS, 2 mM L-glutamine, streptomycin (100 units ml⁻¹), and penicillin (100 μ g ml⁻¹). Cells were differentiated with 1.3% dimethylsulphoxide for 7 days and labelled with 10 mCi ml⁻¹ of myo-[2-³H]inositol (Du Pont NEN) for 20–24 h. ³H-labelled cells were washed with PBS, preincubated for 10 min with increasing concentrations of forskolin in 0.3 mM 3-isobutyl-1-methylxanthine (IBMX), then stimulated with 1 μ M fMLP for 10 min. Agonist-dependent inositol phosphate formation was determined as described²⁶. Shown are the percentage of agonist-stimulated inositol phosphates at increasing concentrations of forskolin. Values are mean values of triplicates; error bar gives the range of triplicates.

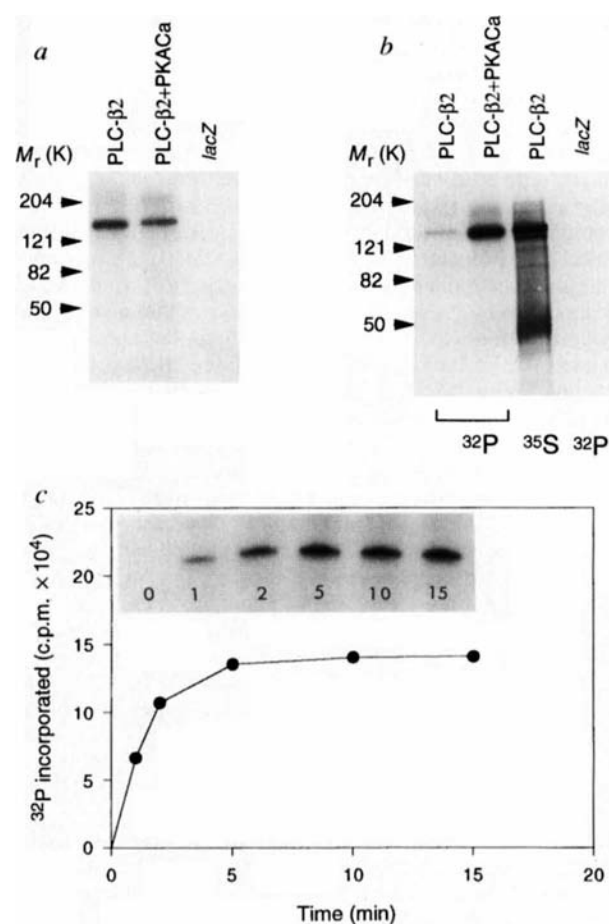
FIG. 3 Phosphorylation of PLC- β 2 by PKA. **a**, Cotransfection of PKA catalytic subunit (Ca) has little effect on the expression of PLC- β 2 protein. COS-7 cells were transiently transfected with a mammalian expressing vector containing the human PLC- β 2 cDNA and PKA Ca. Proteins from cell lysates were separated on a 10% SDS-polyacrylamide gel, immunoblotted with anti-PLC- β 2 antibodies, and detected with enhanced chemiluminescence (ECL, Amersham). There was no immunodetection from mock transfected cells (*lacZ*). In control experiments, even a partial decrease in PLC- β 2 expression has little effect on the activation of PLC- β 2 by G $\beta\gamma$. The threshold for activation of PLC- β 2 by G $\beta\gamma$ was not significantly lower when we transfected different amounts of cDNA (from 0.125 μ g to 1 μ g per well) encoding PLC- β 2 into COS-7 cells. **b**, Metabolic labelling of PLC- β 2 protein in the presence or absence of PKA Ca expression plasmids. COS-7 cells were transfected with cDNAs encoding PLC- β 2 or PLC- β 2 together with a PKA Ca expression plasmid as indicated. Expression plasmid pCIS encoding *lacZ* was used to equalize the total amount of transfected cDNA. Cells were prelabelled with [35 S]methionine or [32 P]orthophosphate and PLC- β 2 protein isolated by immunoprecipitation with anti-PLC- β 2 antibodies (Santa Cruz Biotech). Immunoprecipitates were separated by SDS-PAGE and phosphorylated proteins revealed by autoradiography. **c**, *In vitro* phosphorylation of PLC- β 2 by purified PKA Ca. PLC- β 2 protein was expressed in COS-7 cells and immunopurified by specific PLC- β 2 antibodies. The protein was phosphorylated by purified PKA (Promega) in the presence of [γ - 32 P]ATP at the indicated times. The estimated stoichiometry of PLC- β 2 phosphorylation is >1.7 mol phosphate per mol protein. As the protein was purified by immunoprecipitation, the amount of PLC- β 2 was estimated by the staining intensity with Coomassie brilliant blue.

METHODS. COS-7 cells were transfected by lipofectamine with 2 μ g plasmid DNA per 2.5-cm plate. The PKA Ca plasmid, which encodes the catalytic subunit of PKA (1 μ g), was included where indicated. 48 h after transfection, the cells were labelled with [35 S]methionine (0.5 mCi ml $^{-1}$; trans [35 S]-label; ICN) for 4 h in methionine-free medium or with [32 P]orthophosphate (0.5 mCi ml $^{-1}$; Dupont Biotechnology) for 4 h in phosphate-free medium. After metabolic labelling, cells were lysed in RIPA buffer (50 mM Tris, pH 8.0, 150 mM NaCl, 0.5% deoxycholate, 1% NP-40, 0.1% SDS), supplemented with 50 mM NaF, 1 mM Na $_3$ VO $_4$, aprotinin (2 μ g ml $^{-1}$), leupeptin (2 μ g ml $^{-1}$), pepstatin (1 μ g ml $^{-1}$), 0.25 mM phenylmethanesulphonyl fluoride, 1 mM EDTA. The lysate was centrifuged at 12,000g for 10 min at 4°C. PLC- β 2 protein was immunoprecipitated from 1 ml lysate using 10 μ g affinity-purified rabbit polyclonal antibody directed against a synthetic peptide corresponding to the C-terminal 15 amino-acid residues of PLC- β 2 protein (Santa Cruz Biotech) and protein A-Sepharose CL-4B

Two-dimensional phosphopeptide mapping of 32 P-labelled PLC- β 2 indicated that there are two principal PKA phosphorylation sites. Figure 4a shows the phosphopeptide maps of [32 P]PLC- β 2 isolated from metabolically labelled cells by immunoprecipitation and electrophoresis. In the presence of the PKA catalytic subunit, two additional peptides (arrows) were labelled by [32 P]orthophosphate besides the basal-level phosphopeptides. The same phosphopeptide map was obtained with *in vitro* phosphorylated PLC- β 2 (data not shown). These results indicated that PKA may phosphorylate PLC- β 2 at two main sites that differ from the basal phosphorylation sites. We determined the amino acid phosphorylated by PKA by phosphoamino-acid analysis of isolated 32 P-labelled PLC- β 2, and found that after acid hydrolysis of PLC- β 2 only phosphoserine could be detected, and not phosphothreonine or phosphotyrosine (Fig. 4b).

The amino-acid sequence of PLC- β 2 reveals three possible phosphorylation sites for PKA (Fig. 4c). Peptides corresponding to these three sites were synthesized and purified by reverse-phase HPLC, and then phosphorylated by PKA. Peptide 1 (YNKRQMS 597 RIY) and peptide 3 (GKGSRRKRS 954 LP) are phosphorylated by PKA with a stoichiometry of >0.7 mol phosphate per mol peptide, but peptide 2 (KRRYRT 716 KLS 719 PSP), which contains a threonine was a poor substrate for PKA (Fig. 4d), consistent with serine being the only amino acid that is phosphorylated by PKA in PLC- β 2. These results suggest that Ser 597 and Ser 954 are good candidates for the primary PKA phosphorylation sites in PLC- β 2, but other, less favoured consensus sites in the protein cannot be excluded.

Activation of PLC- β 2 by G-protein $\beta\gamma$ subunits could account for the increase in PTX-sensitive intracellular Ca $^{2+}$ induced by many agonists, including chemoattractants 16 , neuropeptides 22 ,

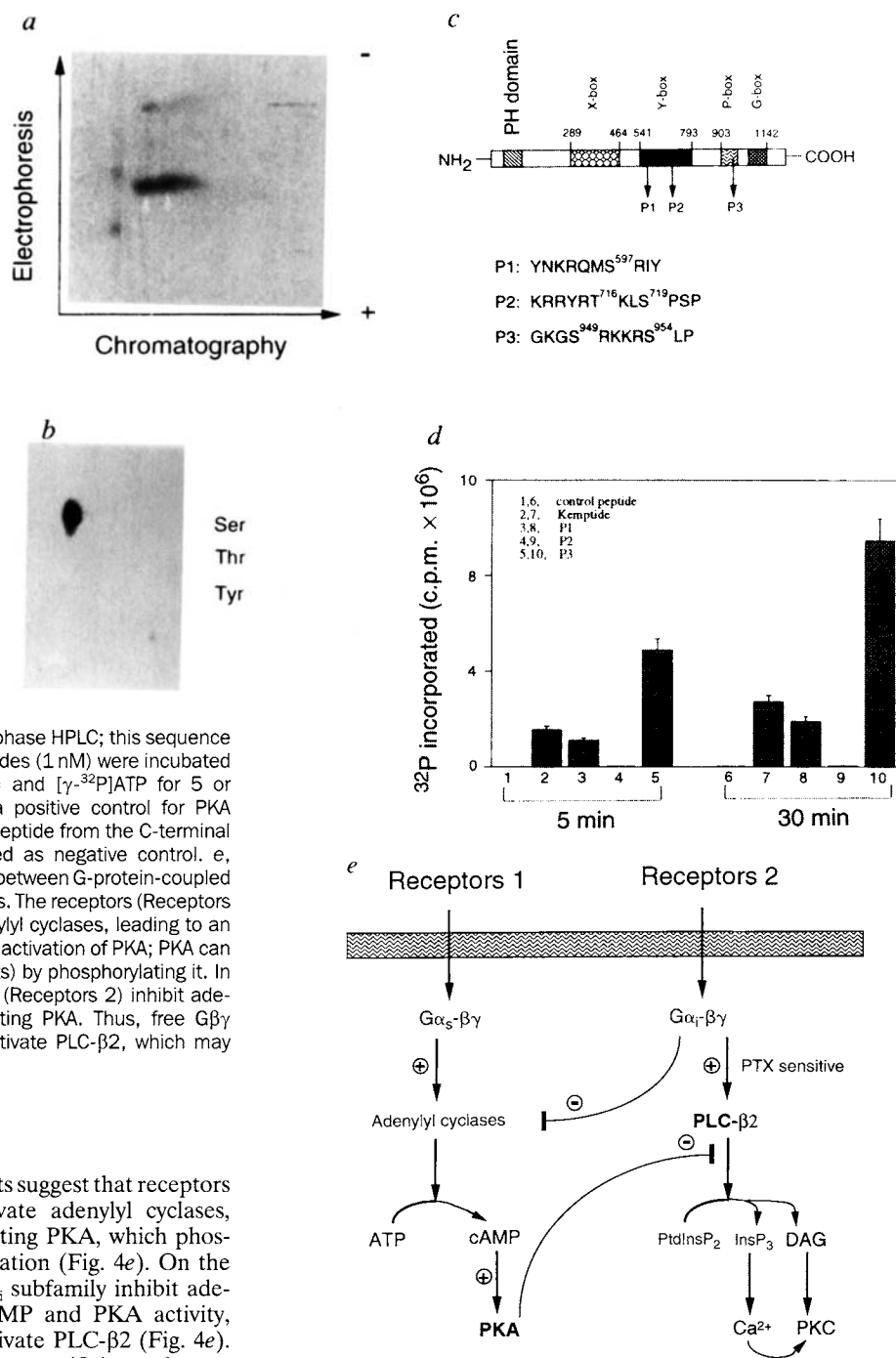


(Sigma). Immunoprecipitates were resolved by SDS-PAGE (7.5 or 10% acrylamide) and autoradiographed.

sphingolipids 23 and other molecules 24 . The region in PLC- β 2 proposed to interact with G $\beta\gamma$ extends from the beginning of the X box to the end of Y box, from residues 289 to 793 (ref. 25). One of the phosphorylated peptides (P1) was found in this region, suggesting that it may be part of the site where phosphorylation modulates the interaction of G $\beta\gamma$ with PLC- β 2. Deletion experiments indicated that the C-terminal region of PLC- β contains a P box (Thr 903 to Gln 1030) and a G box (Lys 1031 to L 1142) 26 . The P box of PLC- β 1 is essential for both its association with the cell membrane and its activation by G α_q , whereas the G box is required for interaction with the G α_q subfamily 26 . One of the phosphopeptides (P3) is in the P-box region of PLC- β 2, so phosphorylation at this site may affect the association of this protein with the membrane. We are investigating whether phosphorylation by PKA affects interaction of PLC- β 2 with G $\beta\gamma$ or with the membrane, or the activity of PLC itself.

There is evidence that phosphorylation is involved in controlling PLC activity. Treatment of intact cells with tumour-promoting phorbol esters (PKC activators) decreases agonist-induced hydrolysis of PtdInsP $_2$ and Ca $^{2+}$ mobilization in some cells 5 . Also, PKA may be involved in regulating the PLC signalling pathway. Treatment of certain cells and tissues with cAMP and agents that stimulate production of cAMP, such as forskolin, theophylline, prostaglandins E1 and E2, prostacyclin and β -adrenergic agents, inhibit phosphoinositide breakdown $^{17-21,27,28}$. Although these cAMP effects are presumably mediated by PKA, it is not clear how PKA modulates the hydrolysis of PtdInsP $_2$. Proteins of the PtdIns signalling pathway proposed as phosphorylation targets include phosphatidylinositol kinase, cell-surface receptors, G proteins and PLC itself $^{28-30}$. We have shown that PKA directly phosphorylates PLC- β 2 and that phosphorylation inhibits activa-

FIG. 4 Location of the major PKA phosphorylation sites in PLC- β 2. **a**, Two-dimensional phosphopeptide map analysis of phosphorylated PLC- β 2. The protein was expressed in COS-7 cells and prelabelled *in vivo* by [32 P]orthophosphate. The PLC- β 2 phosphoprotein was separated by SDS-PAGE, excised from the gel and digested with trypsin ($400 \mu\text{g ml}^{-1}$). After two-dimensional phosphopeptide mapping, phosphopeptides were visualized using a PhosphorImager (Molecular Dynamics). Arrows indicate two major phosphopeptides that are present only in cells cotransfected with PLC- β 2 and PKA catalytic subunit. Similar phosphopeptides were obtained using *in vitro* phosphorylated PLC- β 2 protein. **b**, Phosphoamino-acid analysis of PLC- β 2. 32 P-labelled PLC- β 2 was excised from an SDS-gel, polyacrylamide hydrolysed in 6M HCl, and analysed for 32 P-labelled phosphoamino acids by autoradiography or using a phosphorImager. Positions of standard phosphoamino acids are indicated. **c**, Different domain structures and major putative PKA phosphorylation sites of PLC- β 2 protein. **d**, Phosphorylation of PLC- β 2 peptides by PKA. Peptides P1, P2 and P3, corresponding to three putative PKA phosphorylation sites, were synthesized and purified by reverse-phase HPLC; this sequence was confirmed by mass spectrometry. The peptides (1 nM) were incubated with purified PKA catalytic subunit (Promega) and [γ - 32 P]ATP for 5 or 30 min. Kempptide (LRRASLG) was used as a positive control for PKA phosphorylation reactions²⁷. A 15-amino-acid peptide from the C-terminal end of PLC- β 2 (QDPLIAKADAQESRL) was used as negative control. **e**, Proposed model for the specificity and crosstalk between G-protein-coupled cAMP and inositol phosphate signalling pathways. The receptors (Receptors 1) coupled to G_{α_s} subunits can stimulate adenylyl cyclases, leading to an increase in intracellular cAMP from ATP and the activation of PKA; PKA can inhibit the activation of PLC- β 2 (by $G_{\beta\gamma}$ subunits) by phosphorylating it. In the right-hand pathway, G_{α_i} -coupled receptors (Receptors 2) inhibit adenylyl cyclases, decreasing cAMP and inactivating PKA. Thus, free $G_{\beta\gamma}$ subunits dissociated from G_{α_i} proteins can activate PLC- β 2, which may account for the PTX-sensitive activity of PLC.



tion of PLC- β 2 by $G_{\beta\gamma}$ subunits. Our results suggest that receptors coupling to G-protein α_s subunits activate adenylyl cyclases, increasing intracellular cAMP and activating PKA, which phosphorylates PLC- β 2 and prevents its activation (Fig. 4e). On the other hand, receptors coupled to the G_{α_i} subfamily inhibit adenylyl cyclases, lowering intracellular cAMP and PKA activity, thereby helping free $G_{\beta\gamma}$ subunits to activate PLC- β 2 (Fig. 4e). Such a mechanism would account for the specificity and communication between the cAMP and the phosphatidylinositol signalling pathways.

Received 21 February; accepted 14 May 1996.

- Gilman, A. G. *Rev. Biochem.* **56**, 615–649 (1987).
- Simon, M. I., Strathmann, M. P. & Gautam, N. *Science* **252**, 802–808 (1991).
- Neer, E. J. *Cell* **80**, 249–257 (1995).
- Berridge, M. J. *Nature* **361**, 315–325 (1993).
- Rhee, S. G. & Choi, K. D. *J. Biol. Chem.* **267**, 12393–12396 (1992).
- Smrcka, A. V., Hepler, J. R., Brown, K. O. & Sternweis, P. C. *Science* **251**, 804–807 (1991).
- Taylor, S. J., Chae, H. Z., Rhee, S. G. & Exton, J. H. *Nature* **350**, 516–518 (1991).
- Jhon, D. Y. *et al. J. Biol. Chem.* **268**, 6654–6661 (1993).
- Jiang, H. P., Wu, D. & Simon, M. I. *J. Biol. Chem.* **269**, 7593–7596 (1994).
- Camps, M. *et al. Nature* **360**, 684–688 (1992).
- Katz, A., Wu, D. & Simon, M. I. *Nature* **360**, 686–689 (1992).
- Uhler, M. D. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 1300–1304 (1986).
- Dekker, L. V. & Parker, P. J. *Trends biochem. Sci.* **19**, 73–77 (1994).
- Ryu, S. H. *et al. J. Biol. Chem.* **265**, 17941–17945 (1990).
- Offermanns, S. & Simon, M. I. *J. Biol. Chem.* **270**, 15175–15180 (1995).
- Wu, D., LaRosa, G. J. & Simon, M. I. *Science* **261**, 101–103 (1993).
- Watson, S. P., McConnell, R. T. & Lipetina, E. G. *J. Biol. Chem.* **259**, 13199–13203 (1984).
- Hall, I. P., Donaldson, J. & Hill, S. J. *Br. J. Pharmacol.* **100**, 646–650 (1990).

19. Takenawa, T., Ishitoya, J. & Nagai, Y. *J. Biol. Chem.* **261**, 1092–1098 (1986).
20. Neylon, C. D. & Summers, R. J. *Eur. J. Pharmacol.* **148**, 441–444 (1988).
21. Madison, J. M. & Brown, J. K. *J. clin. Invest.* **82**, 1462–1465 (1988).
22. Perney, T. M. & Miller, R. J. *J. Biol. Chem.* **264**, 7317–7327 (1989).
23. Okajima, F. & Kondo, Y. *J. Biol. Chem.* **270**, 26332–26340 (1995).
24. Snyderman, R., Perlanin, A., Evans, T., Polakis, P. & Didsburg, J. in *ADP-ribosylating Toxin and G Proteins* (eds Moss, J. & Vaughan, M.) 295–323 (Am. Soc. Microbiol., Washington DC, 1990).
25. Wu, D., Katz, A. & Simon, M. I. *Proc. natn. Acad. Sci. U.S.A.* **90**, 5297–5301 (1993).
26. Wu, D., Jiang, H., Katz, A. & Simon, M. I. *J. Biol. Chem.* **268**, 3704–3709 (1993).
27. Anwer, K., Hovington, J. A. & Sanborn, B. M. *Endocrinology* **124**, 2995–3002 (1989).
28. Wen, Y., Anwer, K., Singh, S. P. & Sanborn, B. M. *Endocrinology* **131**, 1377–1382 (1992).
29. Kim, U. H., Kim, J. W. & Rhee, S. G. *J. Biol. Chem.* **264**, 20167–20170 (1989).
30. Richardson, R. M., Ali, H., Tomhave, E. D., Haribabu, B. & Snyderman, R. J. *J. Biol. Chem.* **270**, 27829–27833 (1995).

ACKNOWLEDGEMENTS. We thank M. D. Uhler (University of Michigan) for cDNA clones of the catalytic subunit of the cAMP-dependent protein kinase, P. J. Parker (Imperial Cancer Research Fund, London) for cDNA clones of protein kinase C isoforms, and L. Alex, G. Sanna, L. Brundage, S. Offermanns, B. Yu, A. Aragay, J. Chen, S. Klump and D. Wu for their comments and discussions. This work was supported by a grant from the NIH.

CORRESPONDENCE and requests for materials should be addressed to M.I.S. (e-mail: simonm@starbase1.caltech.edu).

Repression of RNA polymerase III transcription by the retinoblastoma protein

Robert J. White^{*‡}, Didier Trouche^{†‡}, Klaus Martin[†], Stephen P. Jackson^{*} & Tony Kouzarides[†]

Wellcome/CRC Institute, Tennis Court Road, Cambridge CB2 1QR, and ^{*} Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ and [†] Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QP, UK [‡] Present addresses: Institute of Biomedical and Life Sciences, Division of Biochemistry and Molecular Biology, Davidson Building, University of Glasgow, Glasgow G12 8QQ, UK (R.J.W.); URA 1156 CNRS, Institut Gustave Roussy, 39 Rue C. Desmoulins, 94805 Villejuif Cedex, France (D.T.).

TRANSCRIPTION by RNA polymerase (pol) III is under cell-cycle control, being higher in S and G2 than in G0 and early G1 phases¹⁻⁴. Many transformed cell types have elevated pol III activity⁴⁻⁹, presumably to sustain sufficient protein synthesis for unrestrained growth. The retinoblastoma tumour-suppressor protein (Rb) restricts cellular proliferation, and is often found mutated in transformed cells^{10,11}. Here we demonstrate that Rb can repress the level of transcription from pol III templates both *in vitro* and *in vivo*. Analysis of Rb-deficient SAOS2 cells and primary fibroblasts from *Rb*^{-/-} mice demonstrates elevated levels of pol III activity in the absence of functional Rb protein. Rb-induced repression of pol III activity is alleviated by mutations in

the Rb pocket domain that occur naturally in tumours, and by viral transforming proteins that bind and inactivate Rb. These results implicate repression of pol III transcription as a mechanism for Rb-induced growth arrest, and suggest that restraining protein biosynthesis may be important in the prevention of tumour development.

There is a degree of sequence similarity between Rb and two general transcription factors, TATA-binding protein (TBP) and TFIIB¹². This has led to the proposal that Rb may repress E2F-regulated pol II transcription, at least partly, by mimicking the structure of these two factors^{11,13}. Because TBP- and TFIIB-like factors are required for the transcription of pol III as well as pol II promoters^{14,15}, we investigated whether Rb can repress transcription

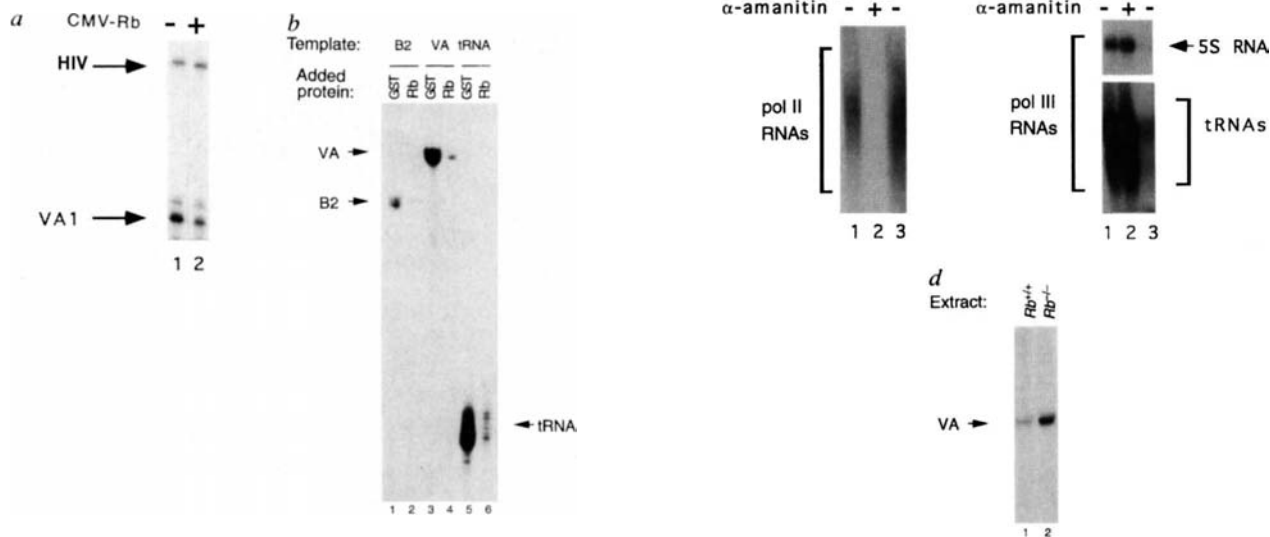


FIG. 1 Rb represses transcription of pol III templates. *a*, U2OS cells ($\sim 10^6$) were transfected overnight by calcium phosphate coprecipitation with 4 μ g pVA_i (ref. 21) and 4 μ g HIV-CAT²² without (lane 1) or with (lane 2) 1 μ g CMV-Rb²³. The amount of CMV promoter was kept constant by using the parental CMV vector. Total RNA was extracted 24 h later and analysed by primer extension. *b*, Pol III transcription of B2 (lanes 1 and 2), VA_i (lanes 3 and 4) and tRNA (lanes 5 and 6) genes (500 ng) reconstituted using 2 μ l each of 0.38 M TFIIB²⁴ (0.90 μ g), 0.48 M TFIIB²⁴ (0.22 μ g) (**OK?**) and PC-C²¹ (2.6 μ g) fractions preincubated with 250 ng GST (odd-numbered lanes) or GST-Rb (residues 379-928) (even-numbered lanes). All reactions contained 1 μ g ml⁻¹ α -amanitin.

METHODS. *a*, Total RNA (20 μ g), extracted using an ultraspec-II system (Biotecx), was analysed by primer extension using VA_i-specific (5'-GGGGTTCGAACCCCGTCGTC-3') and CAT-specific (5'-GGAGCTAAGGAAGC-TAAAT-3') labelled primers. Primer extension was performed with 50 μ g ml⁻¹ actinomycin D as described previously²², except that the samples were incubated for 10 min at 80 °C before annealing. *b*, Templates, protein fractions and transcription conditions have been described^{21,24}. GST-Rb was prepared as described previously¹².

FIG. 2 Cells lacking functional Rb have elevated levels of pol III activity. *a*, U2OS (lanes 1 and 2) and SAOS2 (lanes 3 and 4) cells ($\sim 10^6$) were transfected with pVA_i (4 μ g) and HIV-CAT²² (4 μ g). Total RNA was analysed as in Fig. 1a. *b*, Transcription of pVA_i (250 ng) using 3.6 μ g whole-cell extract of U2OS (lane 1) or SAOS2 (lane 2) cells. *c*, Nuclei (10^7) of embryonic fibroblasts from *Rb*^{-/-} (lanes 1 and 2) or wild-type (lane 3) mice²⁵ were used in a transcriptional run-on experiment in the presence (lane 2) or absence (lanes 1 and 3) of 0.5 μ g ml⁻¹ α -amanitin. Labelled RNA products were separated on 10% polyacrylamide gel and detected by autoradiography²⁶. Left, the top of the gel; it represents RNAs larger than 500 bases. Right, the bottom of the gel; the 5S rRNA co-migrates with a 126-bp DNA marker, and the tRNA migrates at 80-110 bases. *d*, Transcription of pVA_i (250 ng) using 24 μ g whole-cell extract of *Rb*^{-/-} (lane 1) or *Rb*^{+/+} (lane 2) mouse embryo fibroblasts.

METHODS. *a*, As Fig. 1a. *b*, *d*, Extracts were prepared as described previously²¹. *c*, Primary embryonic fibroblasts from *Rb*^{-/-} and *Rb*^{+/+} mice²⁵ were grown for a similar number of passages in DMEM supplemented with 20% FCS. Nuclear extraction, run-on reactions and electrophoresis of labelled RNAs were performed as described previously²⁶.

of a pol III template, the VA₁ gene of adenovirus. VA₁ gene transcription is indeed repressed by Rb in U2OS cells (Fig. 1a). In contrast, the pol II-transcribed human immunodeficiency virus (HIV) promoter is not affected, indicating that Rb can regulate a pol III template in a specific manner.

To determine whether the effect of Rb upon VA₁ transcription

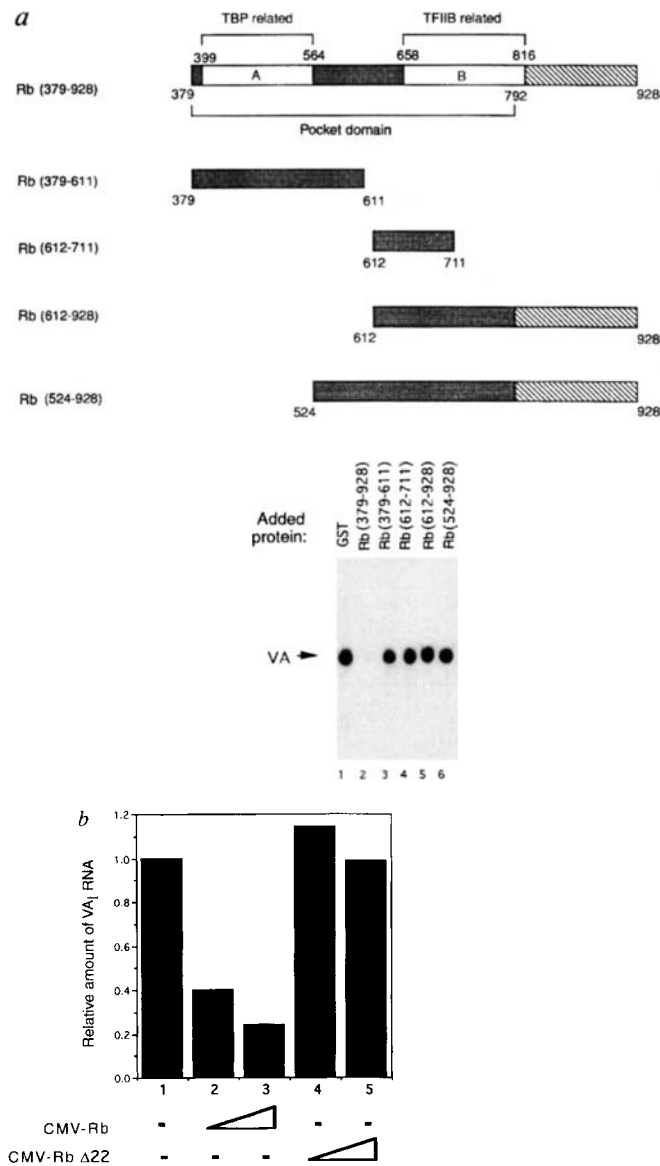


FIG. 3 Mutations in the Rb pocket domain prevent repression of pol III transcription. **a**, Transcription of pVA₁ (250 ng) using 1 μ l each of 0.38 M TFIIB (0.45 μ g) and 0.48 M TFIIB²⁴ (0.11 μ g) fractions, and 2 μ l (1.2 μ g) of TFIIC and pol III-containing CHeP-1.0 fraction²¹ preincubated (15 min, 30 °C) with 250 ng GST (lane 1), GST-Rb (379-928) (lane 2), GST-Rb (379-611) (lane 3), GST-Rb (612-711) (lane 4), GST-Rb (612-928) (lane 5), or GST-Rb (524-928) (lane 6). Bracketed numbers indicate the Rb residues included in a fusion protein. **b**, U2OS cells ($\sim 10^6$) were transfected with pVA₁ (4 μ g) and 0.5 or 2 μ g of CMV-Rb (lanes 2 and 3, respectively), or 0.5 or 2 μ g of CMV-RbΔ22 (ref. 23) (lanes 4 and 5, respectively). The amount of CMV promoter was kept constant using parental CMV vector. Total RNA was analysed as in Fig. 1. VA₁ RNA was quantified relative to endogenous β -actin RNA by phosphorimager analysis. The amount of VA₁ RNA in the absence of Rb was normalized to a value of 1. **c**, Transcription of pVA₁ (250 ng) using 1 μ l each of 0.38 M TFIIB (0.45 μ g) and 0.48 M TFIIB²⁴ (0.11 μ g) and 2 μ l of CHeP-1.0 (ref. 21) (1.2 μ g) fractions preincubated (15 min, 30 °C) with 250 ng GST (lanes 1 and 4), GST-Rb (379-928) (lanes 2 and 5), GST-Δ21 mutant Rb²⁷ (lane 3), and GST-Rb (379-928) with a single substitution at residue 706 of cysteine to phenylalanine²⁷ (lane 6).

is direct, we tested whether this template can be repressed by Rb *in vitro*. A glutathione *S*-transferase (GST)-Rb fusion protein containing residues 379 to 928 of Rb was found to repress VA₁ transcription that had been reconstituted using partly purified factors (Fig. 1b). Rb repression is not restricted to VA₁; every pol III template we have tested is responsive to Rb repression, including various transfer RNA, U6 small nuclear RNA, B2 EBER2 and 5S ribosomal RNA genes (Fig. 1b and data not shown). We conclude that Rb can function as a general repressor of pol III transcription.

If Rb is important for pol III inhibition under physiological conditions, there should be elevated expression of class III genes in Rb-deficient cells. We therefore compared the pol III activity of two osteosarcoma cell lines: U2OS, which contains functional Rb, and SAOS2, which does not. Transcription of the VA₁ gene was found to be much more active in Rb-deficient SAOS2 cells than in Rb-positive U2OS cells, both *in vivo* (Fig. 2a) and *in vitro* (Fig. 2b). To confirm that elevated pol III activity was indeed caused by deficiency in Rb, we assessed the level of pol III activity in primary fibroblasts from wild-type or Rb-deficient mice. Nuclear run-on assays were performed using Rb^{+/+} and Rb^{-/-} primary fibroblasts, and the labelled pol II and pol III RNAs (distinguished by their α -amanitin sensitivity) were detected following electrophoresis. These assays showed that the synthesis of transfer RNA and 5S ribosomal RNA by pol III is substantially elevated in Rb^{-/-} fibroblasts relative to Rb^{+/+} fibroblasts (Fig. 2c). In contrast, the total level of pol II transcription is not increased by deletion of the Rb gene (Fig. 2c). *In vitro* transcription assays using extracts of these primary fibroblasts provided independent confirmation that the pol III transcription apparatus is more active in the absence of Rb (Fig. 2d). The difference in pol III activity between Rb-positive and Rb-negative cells was seen under a variety of assay conditions, and with several independently prepared sets of extracts. These data demonstrate that endogenous Rb is involved in repressing pol III transcription, and corroborate the results in Fig. 1b showing that Rb can repress transcription of a variety of pol III genes.

The pocket domain of Rb must be intact for the binding and repression of E2F and the growth-suppressive activity of Rb. Many tumours contain mutations in the Rb pocket that inactivate these functions¹⁶. Deletion analysis of Rb shows that an intact pocket is also required for the inhibition of pol III transcription by Rb (Fig. 3a). In addition, naturally occurring deletions of the Rb pocket, found in tumours (exon 22 and 21 deletions¹⁶) disrupt the

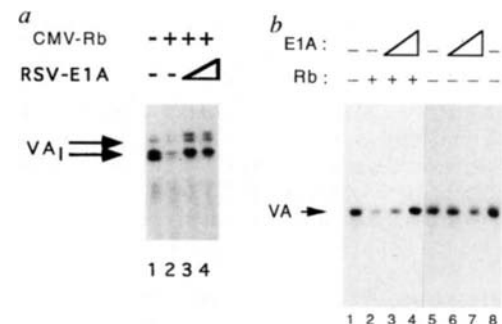


FIG. 4 Adenovirus E1A can overcome repression of pol III transcription by Rb. **a**, Transient transfection of U2OS cells ($\sim 10^6$) with pVA₁ (4 μ g), 2 μ g CMV-Rb (lanes 2-4) and 2.5 or 5 μ g RSV-E1A (lanes 3 and 4, respectively). The amounts of CMV and RSV promoters were kept constant using parental CMV and RSV vectors. Total RNA was analysed as in Fig. 1. **b**, Transcription of pVA₁ (250 ng) using 1 μ l each of 0.38 M TFIIB (0.45 μ g) and 0.48 M TFIIB²⁴ (0.11 μ g) and 2 μ l CHeP-1.0 (ref. 21) (1.2 μ g) fractions preincubated (15 min, 30 °C) with 250 ng GST (lanes 1 and 5-8) or GST-Rb (379-928) (lanes 2-4), and 50 ng (lanes 3 and 6) or 200 ng (lanes 4 and 7) of GST-E1A.

METHODS. **a**, RSV-E1A contains the 12S coding region of the E1A gene fused to the RSV promoter. **b**, GST-E1A consists of the 289 amino-acid E1A protein fused to GST.

ability of Rb to repress pol III transcription *in vitro* and *in vivo* (Fig. 3b,c). Furthermore, an inactivating point mutation found in a small cell lung carcinoma which causes substitution of residue 706 from a cysteine to a phenylalanine¹⁷, completely abolishes the ability of Rb to repress VA₁ transcription (Fig. 3c). These results demonstrate that the ability of Rb to function as a tumour suppressor correlates with its ability to repress pol III transcription.

Several viral oncoproteins, such as adenovirus E1A and SV40 large T antigen, can bind and inactivate Rb¹⁰. VA₁ transcription has been shown to be activated by both E1A and T antigen in transfected cells¹⁸, and by E1A *in vitro*¹⁹. Repression of VA₁ by Rb can be overcome *in vivo* by the 243 amino-acid E1A protein (Fig. 4a). This is likely to be a direct effect, as E1A also relieves Rb repression of transcription reconstituted with partly purified factors (Fig. 4b). SV40 large T antigen can also overcome the inhibition of VA₁ expression by Rb (data not shown). Although E1A and T antigen regulate transcription by multiple mechanisms, these results suggest that one way in which oncoproteins activate pol III transcription is by relieving the repressive effects of Rb.

This demonstration that Rb can repress pol III transcription, along with data showing that Rb can repress pol I (ref. 20) and pol II (ref. 10, 11) transcription, indicates that Rb is the first characterized example of a factor that can negatively regulate the activity of all three classes of nuclear RNA polymerases. Furthermore, our data indicate that the tumour suppressor functions of Rb correlate not only with the suppression of S phase (by silencing E2F) but also with the suppression of pol III transcription. This raises the possibility that Rb-mediated growth control may involve the restraint of protein biosynthesis. □

Received 1 August 1995; accepted 16 May 1996.

1. Johnson, L. F., Abelson, H. T. & Green, H. *Cell* **1**, 95–100 (1974).
2. Mauck, J. C. & Green, H. *Cell* **3**, 171–177 (1974).
3. White, R. J., Gottlieb, T. M., Downes, C. S. & Jackson, S. P. *Molec. cell. Biol.* **15**, 6653–6662 (1995).
4. White, R. J. *RNA Polymerase III Transcription* (R. G. Landes, Austin TX, 1994).
5. Scott, M. R. D., Westphal, K.-H. & Rigby, P. W. J. *Cell* **34**, 557–567 (1983).
6. Majello, B. et al. *Nature* **314**, 457–459 (1985).
7. Singh, K., Carey, M., Saragosti, S. & Botchan, M. *Nature* **314**, 553–556 (1985).
8. Carey, M. F., Singh, K., Botchan, M. & Cozzarelli, N. R. *Molec. cell. Biol.* **6**, 3068–3076 (1986).
9. White, R. J., Stott, D. & Rigby, P. W. J. *EMBO J.* **9**, 3713–3721 (1990).
10. Weinberg, R. A. *Cell* **81**, 323–330 (1995).
11. Kouzarides, T. *Cancer Biol.* **6**, 91–98 (1995).
12. Hagemeier, C., Bannister, A. J., Cook, A. & Kouzarides, T. *Proc. natn. Acad. Sci. U.S.A.* **90**, 1580–1584 (1993).
13. Hagemeier, C., Cook, A. & Kouzarides, T. *Nucleic Acids Res.* **21**, 4998–5004 (1993).
14. Rigby, P. W. J. *Cell* **72**, 7–10 (1993).
15. Hernandez, N. *Genes Dev.* **7**, 1291–1308 (1993).
16. Horowitz, J. M. et al. *Proc. natn. Acad. Sci. U.S.A.* **87**, 2775–2779 (1990).
17. Kaye, F. J., Kratzke, R. A., Gerster, J. L. & Horowitz, J. M. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6922–6926 (1990).
18. Loeken, M., Bikel, I., Livingston, D. M. & Brady, J. *Cell* **55**, 1171–1177 (1988).
19. Patel, G. & Jones, N. C. *Nucleic Acids Res.* **18**, 2909–2915 (1990).
20. Cavanaugh, A. H. et al. *Nature* **374**, 177–180 (1995).
21. White, R. J., Gottlieb, T. M., Downes, C. S. & Jackson, S. P. *Molec. cell. Biol.* **15**, 1983–1992 (1995).
22. Trouche, D., Grigoriev, M., Robin, P. & Harel-Bellan, A. *Biochem. biophys. Res. Commun.* **196**, 611–618 (1993).
23. Qin, X.-Q., Chittenden, T., Livingston, D. M. & Kaelin, W. G. Jr *Genes Dev.* **6**, 953–964 (1992).
24. Lobo, S. M., Tanaka, M., Sullivan, M. L. & Hernandez, N. *Cell* **71**, 1029–1040 (1992).
25. Jacks, T. et al. *Nature* **359**, 295–300 (1992).
26. Marzluff, W. F. & Huang, R. C. C. in *Transcription and Translation: a Practical Approach* (eds Hames, B. D. & Higgins, S. J.) 89–129 (IRL, Oxford, 1984).
27. Kaelin, W. G. Jr et al. *Cell* **64**, 521–532 (1991).

ACKNOWLEDGEMENTS. R.J.W. and D.T. contributed equally to this work. We thank R. Herrerra, T. Jacks and M. Carlton for mouse embryo fibroblasts; W. G. Kaelin and A. J. Bannister for GST–Rb constructs; and H. Land for RSV–E1A. This work was funded by groups from the Cancer research campaign and the Medical Research Council.

CORRESPONDENCE and requests for materials should be addressed to T.K. (e-mail: tk106@ mole.bio.cam.ac.uk) or R.J.W. (rwhite@udcf.gla.ac.uk).

A helical arch allowing single-stranded DNA to thread through T5 5'-exonuclease

T. A. Ceska, J. R. Sayers*, G. Stier & D. Suck

EMBL, Structural Biology Programme, Meyerhofstrasse 1, 69117 Heidelberg, Germany

* University of Sheffield, Krebs Institute, Department of Medicine and Pharmacology, Section of Molecular Medicine, Royal Hallamshire Hospital, Sheffield S10 2JF, UK

THE 5'-exonucleases are enzymes that are essential for DNA replication and repair¹. As well as their exonucleolytic action, removing nucleotides from the 5'-end of nucleic acid molecules such as Okazaki fragments², many 5'-3'-exonucleases have been shown to possess endonucleolytic activities^{3,4}. T5 5'-3'-exonuclease shares many similarities with the amino termini of eubacterial DNA polymerases⁵, although, unlike eubacteria, phages such as T5, T4 and T7 express polymerase and 5'-exonuclease proteins from separate genes. Here we report the 2.5-Å crystal structure of the phage T5 5'-exonuclease, which reveals a helical arch for binding DNA. We propose a model consistent with a threading mechanism in which single-stranded DNA could slide through the arch, which is formed by two helices, one containing positively charged, and the other hydrophobic, residues. The active site is at the base of the arch, and contains two metal-binding sites.

The T5 D15 5'-exonuclease is a protein of relative molecular mass 33,000 (*M_r* 33K). Its substrates include single- and double-stranded linear DNAs, which are degraded to short oligomers⁶ and flap structures, which are first cut at the junction of the single- and double-stranded DNA. The purification and crystallization

conditions for T5 5'-exonuclease have been reported previously⁷. The structure determination was hampered by variability of the space group of native crystals, and the dearth of good heavy-atom derivatives. Only mercury compounds were found to derivatize the crystals, and mercury binding was accompanied by non-isomorphous changes. The native non-isomorphism was overcome by preparing a selenomethionine analogue of the nuclease, which crystallized under similar conditions to the native protein, and collecting multiwavelength anomalous dispersion data at 2.5 Å (Table 1).

The structure of T5 D15 5'-exonuclease is shown in Fig. 1a. The central β-sheet consists of 6 strands, and has a cluster of several α-helices on either side. We traced residues 20 to 291, including all of the conserved residues identified by sequence analysis^{5,8}. The N-terminal 18 residues are disordered in the T5 5'-exonuclease, but they are remote from the active site. The first 17 amino acids can be removed by *in vitro* proteolysis without altering enzymatic activity (J.R.S. and D. Evans, unpublished results). The reported structure of the 5'-exonuclease domain of *Taq* polymerase⁹ has extensive regions of disorder (up to 30% of the domain), but the ordered regions are roughly similar to our T5 exonuclease. However, we have observed some interesting new features in our structure. The active site in the *Taq* 5'-exonuclease was described as a deep cleft, but there is no cleft in our structure, only the concave surface in the front of the protein which includes the active-site area. A triangular structure at the back of the nuclease is present, and comprises three α-helices that form an arch. There are positively charged residues on helix 4, and hydrophobic residues on helix 5. The rear of the helical-arch motif carries several negatively charged amino acids, which could facilitate product release after cleavage of the phosphodiester bond. Sequence alignments and secondary structure predictions for *Taq* and other DNA polymerases indicate that there is a preference for α-helical structures in a region corresponding to the helical arch of the T5 nuclease¹⁰. Thus some form of arch or loop may exist in homologous 5'-exonucleases. The side chains of two absolutely

FIG. 1 a, The structure of the T5 5'-exonuclease showing residues 20–290 (drawing produced using the program O²²). The residues presumed to be involved in single-stranded DNA binding (helix 4, K83, R86, K89 and R93; helix 5, F104, F105; helix 1, F32) and catalysis are shown. The active site contains two metal-binding sites. The divalent cations required for catalysis are coordinated by aspartic acid residues. Data collected from crystals soaked in 25 mM MnCl₂ reveal the location of the metals, and the magenta-coloured difference electron density is contoured at 4 σ . The Mn site labelled 1 is also the site for Zn binding. Helices 3 and 4 contain arginine and lysine residues along the interior of the helices. This is the DNA-binding motif, which is disordered in the 5'-exonuclease domain of *Taq* polymerase⁹. The basic nature of helix 4 implicates it in DNA binding. The other side of the helical arch possesses a hydrophobic patch of 3 phenylalanine residues that might interact with the bases of single-stranded DNA. b, A stereo view showing the positions of the 14 conserved residues^{5,8} in the 5'-exonuclease family. The aspartic acid residues (D26, D128, D131 and D153) coordinate Mn1, and aspartic acid residues (D153, D155, D201 and D204) coordinate Mn2. There are additional acidic residues not strictly conserved, that also appear to coordinate the metals that are not shown. The site labelled Mn1 corresponds to site III, and Mn2 corresponds to site II in the 5'-exonuclease domain of *Taq* polymerase. There are two fully conserved basic residues, K83 and R86, on the inside of helix 4. There are also two other basic residues that point in the same direction (K89 and R93). The conserved residue Y82 is located at the base of the arch.

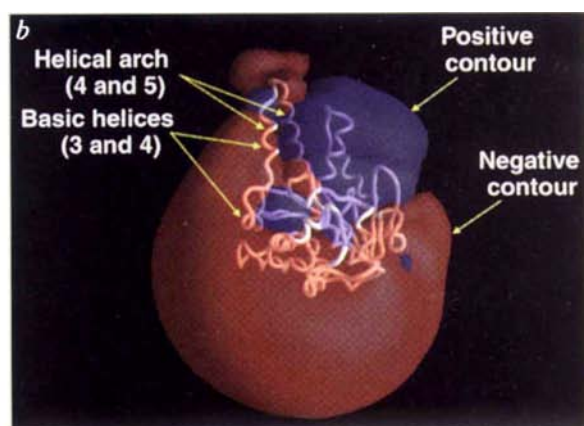
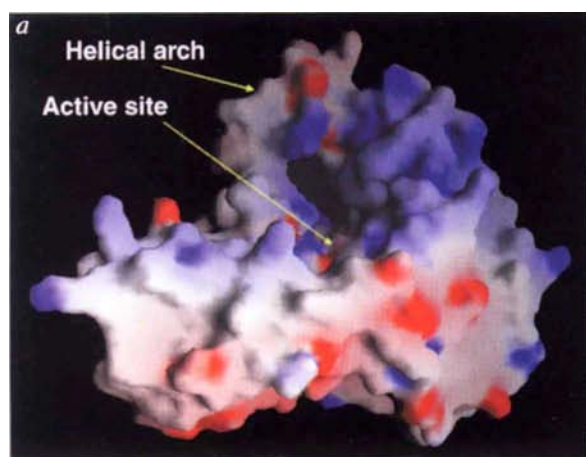
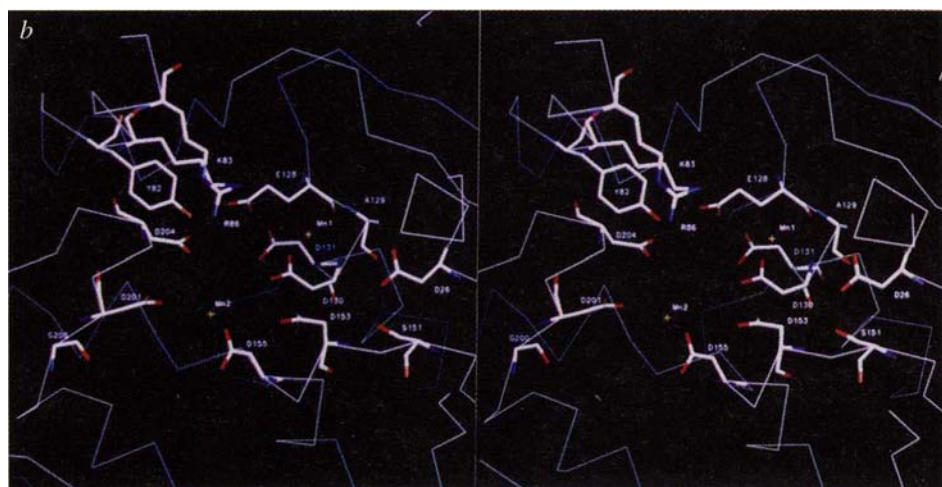
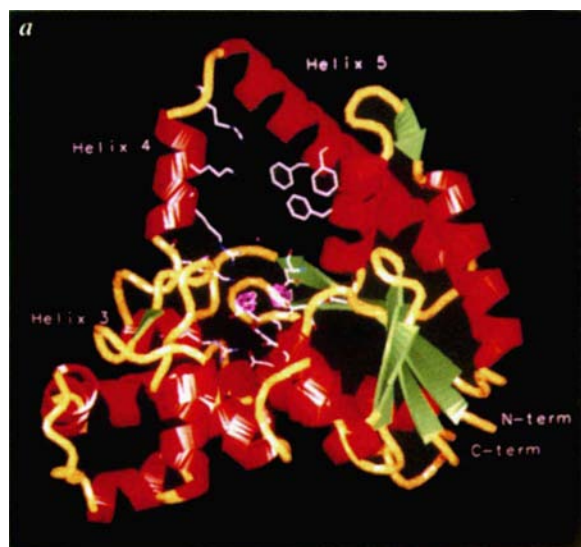


FIG. 2 a, T5 5'-exonuclease, showing the charge distribution mapped to the surface of the molecule. The model created with GRASP²⁵, includes the two Mn²⁺ atoms bound in the active site. The regions of basic potential ($>9 k_B T/e$) are shown in blue, and acidic ($<-11 k_B T/e$) in red. The concave surface of the molecule leading up to the helical arch is shown in this view, and is the location of the active site. The basic nature of the helical arch should strongly attract DNA. b, An electrostatic potential surface contour

plot, created with GRASP, showing the bipolar distribution of charge of T5 5'-exonuclease viewed from the side. The contour level was set at $\pm 2 k_B T/e$. The back of the helical arch has a negative potential (red) (there are several glutamic acid residues on this surface), and the front of the helical arch has a positive potential (blue). From this view it can be seen that DNA can only approach the active site from one side.

conserved basic residues (Lys 83 and Arg 86) are in the arch, and the absolutely conserved tyrosine residue (Tyr 82) also forms part of the floor of the arch. The conserved residues in the 5'-exonuclease family are shown in Fig. 1b.

The electrostatic charge potential of the molecule mapped to the surface is depicted in Fig. 2a, showing the charge distribution on the molecule. Helices 4 and 5, which form the arch, show bipolar and amphipathic character, respectively. The positively charged residues (Lys 83, Arg 86, Lys 89 and Arg 93) are all directed towards the interior of the arch, presumably to allow interaction with the DNA-phosphate backbone. The opposite face of this helix carries two negatively charged residues (Asp 87 and Glu 88). In addition to the basic residues along one side of the arch, there is a cluster of phenylalanine residues on the other side of the arch pointing into the central region, that are candidates for base binding, as has been observed in bacteriophage fd gene 5 protein-DNA complexes¹¹, and in the binding of nucleotides to the 3'-5'-exonuclease site of the Klenow fragment¹². The electrostatic potential contours around the molecule suggest the direction of DNA approach (Fig. 2b).

Some proteins have previously been shown to have holes that accommodate DNA, for example topoisomerase^{13,14} and DNA gyrase B protein¹⁵ among others, but in these cases the holes accommodate double-stranded DNA. The helical arch motif forms a hole big enough to accommodate single-stranded but not double-stranded DNA. The active site is located in front of the arch, and comprises a cluster of aspartic acid residues that complex two divalent cations. The shape of the platform of the enzyme is slightly concave, presumably to complement the shape of the DNA, and this surface contains several lysine and arginine residues.

Human DNase IV¹⁶, mouse FEN-1 nuclease⁴, and calf 5'-3'-exonuclease¹⁷ are also able to act as structure-specific nucleases in a manner analogous to the N terminus of the DNA polymerase I of *Escherichia coli*, as well as having 5'-exonuclease activity. The

enzymes isolated from eukaryotes and prokaryotes share some sequence conservation as well as having similar activities. In flap structures, once the displaced single-stranded DNA becomes threaded, the enzyme slides along until it reaches the duplex part of the structure. The duplex binds to the enzyme and cleavage subsequently occurs. The products of homopolymer digestion catalysed by T5 5'-exonuclease are short oligonucleotides, mainly trimers, tetramers and pentamers⁹. The sliding of the single-stranded DNA through the arch may explain why the reaction products are not unique. Closed circular DNA, with or without mismatches, is not cleaved by this enzyme, showing that a free 5'-end in the substrate is required for nucleolytic activity¹⁸.

It has been shown that a 5'-exonuclease can become trapped on the displaced 5'-strand of a flap structure¹⁷. We show that the T5 5'-exonuclease can cleave flap structures endonucleolytically (Fig. 3a), as well as having an exonuclease function. We modelled a strand-displaced duplex DNA substrate into the helical arch of T5 5'-exonuclease, as shown in Fig. 3b. Examination of the model reveals complementary contacts between conserved basic residues and the DNA backbone, suggesting how T5 5'-exonuclease can use a threading mechanism to translocate DNA to the cleavage site^{3,17}.

A two-metal-ion mechanism, similar to that for the 3'-5'-exonuclease domain of the Klenow fragment¹², may also be relevant for the 5'-exonucleases⁹. We have collected data in which the crystals have been soaked in 25 mM MnCl₂ and observe two metal sites, separated by 8.1 Å, which suggests that these are the positions of the divalent cations required for activity. The aspartic acid residues, which are conserved, and the environment of the divalent cation-binding sites are similar to those in *Taq* 5'-exonuclease⁹, although the distance reported between the manganese sites was 10 Å in *Taq* 5'-exonuclease. We have analysed a data set in which a crystal was soaked in both 25 mM MnCl₂ and 2 mM ZnCl₂. We observed only one significant peak in the difference map, in a site identical to one of the Mn sites (Mn1). This contrasts with the *Taq* 5'-exonuclease structure for which a third metal-binding

TABLE 1 Summary of the data collection and phasing statistics

Wavelength (Å)	Resolution (Å)	Number of observations	Unique reflections	Completeness (%)	Bijvoet completeness (%)	R_{sym}^* (%)	Phasing power†‡ (iso/ano)	f'/f'' §
0.9793 Å1 (peak)	2.5	140,024	26,634	96.4	96.9	3.8	2.29/2.45	-4.5/5.8
0.9795 Å2 (inflection)	2.5	140,054	26,640	96.4	96.9	3.8	2.88/2.02	-8.0/2.0
1.035 Å3 (remote)	2.5	139,804	26,671	96.8	96.9	3.7		
0.87 Å4 (remote)	2.5	67,552	26,463	95.2	81.4	3.7	-1.20	-3.0
Overall figure of merit 0.779	Refinement statistics							
	Resolution:			6.0–2.5 Å		Total number of atoms:		
	Number of reflections:			25,534		Water molecules:		
	R.m.s. bond length:			0.011 Å				
	R.m.s. bond angle:			1.7 degrees				
	Crystallographic R-factor (%):			22.7				
	Free-R factor (%):			33.9 for 2,553 reflections				

Protein containing SeMet was expressed in a Met auxotroph strain (LE 392), and SeMet was added to the cell culture after overnight starvation of Met and induced. The protein was purified as previously described⁷, and crystals were obtained under conditions similar to those for the unlabelled protein. The multiwavelength anomalous dispersion (MAD) data were collected at the ESRF in Grenoble on beamline BL19 at 4 wavelengths. The space group for the crystal was determined to be $P4_3$, the most common observed for these crystals. The intensities were integrated with DENZO¹⁹, and the Patterson was solved using the program VERIFY (S. Roderick, unpublished.) PHASES²⁰ was used to calculate an interpretable map. DEMON²¹ was used to get an averaged dimer from an initially determined non-crystallographic symmetry (n.c.s.) operator. The initial model was built with O²² into this averaged map, and after an additional round of n.c.s. improvement and DEMON averaging cycle, a monomer mask was determined from the partly built model. DM²³ was then used to solvent flatten and average the map in 9 stages, with an increasing amount of the model included at each stage. The molecule was refined using XPLOR²⁴, with alternating cycles of refinement and manual rebuilding. The refinement was complicated by the pseudo- $P4_322$ symmetry which exists in the crystal. The experimental phase information and n.c.s. were used first as strong restraints and then weak restraints in the refinement. Residues 20–290 are visible in the map for the first molecule, and 20–291 in the second. The N-terminal methionine (residue 1) is removed in *E. coli*. There are no non-glycine residues in unfavourable regions of a Ramachandran plot.

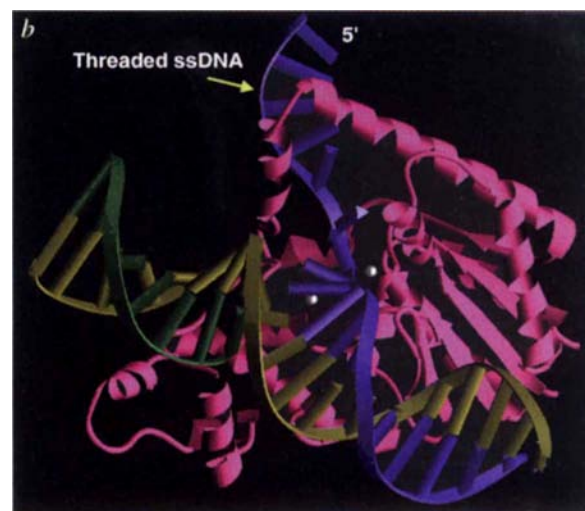
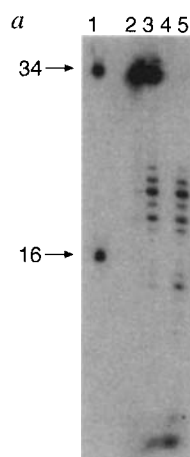
* $R_{\text{sym}} = \sum |I - \langle I \rangle| / \sum I$, where I is the observed intensity, and $\langle I \rangle$ is the average intensity for multiple measurements.

† Phasing power, F_H/E , where F_H is the calculated heavy-atom structure factor, and $E = \text{r.m.s.}(F_H)/\text{r.m.s.}(\text{lack of closure})$.

‡ Resolution threshold was set to 2.8 Å in PHASES.

§ f' and f'' as input into PHASES.

FIG. 3 a, An autoradiogram of the initial products of digestion showing that T5 5'-exonuclease has endonucleolytic activity. Lane 1, 5'-end ^{32}P -labelled oligonucleotide markers of 34 and 16 nucleotides. Unlabelled template and adjacent strands were present at slight molar excess over the 34-mer flap strand (0.2 pmol). The sequence for the flap structure used was as in ref. 4. The substrates were incubated for 30 min at 37 °C in 12 μl of 25 mM glycine/KOH, pH 9.3, 5 mM MgCl_2 and 1 mM DTT with different concentrations of T5 exonuclease. Lane 2, 0.003 pmol enzyme; lane 3, 0.03 pmol enzyme; lane 4, 3.0 pmol enzyme; lane 5, 0.3 pmol enzyme. Reaction products were electrophoresed on a 15% polyacrylamide gel. The initial major products of digestion are 19 and 21 nucleotides in length, consistent with cleavage at the bifurcation. **b**, A conceptual model of how a flap structure could bind to the T5 5'-exonuclease (produced with RIBBONS²⁶). Based upon previous experimental results (refs 3, 7 and our own unpublished observations) we have made a model of the single-stranded flap DNA (ssDNA, blue) threaded through the helical arch. The placement of the DNA was chosen with the electrostatic potential surface contour as a guide. The flap DNA structure was docked on the protein, minimizing clashes between the two molecules. The model was not energy minimized.



The two metal ions are shown as silver spheres. The precise position of the double-stranded DNA docked to the enzyme needs to be determined experimentally.

site occupied by zinc was observed. Thus the distance between the two metal-binding sites we observe is much greater than the usual 4 Å observed in nucleases with a putative two-metal-ion mechanism, and it is likely that the details of the 5'-exonuclease mechanism will be different.

A more complete understanding of the mechanism of the enzyme will clearly require co-crystallization with DNA. Our structure will assist in the devising of the site-directed mutagenesis experiments required to elucidate the mode of action for this member of a biochemically important class of enzymes. □

Received 27 February; accepted 21 May 1996.

- Kornberg, A. & Baker, T. *DNA Replication* 2nd edn (W. H. Freeman, New York, 1992).
- Okazaki, R., Anisawa, M. & Sugino, A. *Proc. natn. Acad. Sci. U.S.A.* **68**, 2954–2957 (1971).
- Lyamichev, V., Brow, M. A. D. & Dahlberg, J. E. *Science* **260**, 778–783 (1993).
- Harrington, J. J. & Lieber, M. R. *EMBO J.* **13**, 1235–1246 (1994).
- Gutman, P. D. & Minton, K. W. *Nucleic Acids Res.* **21**, 4406–4407 (1993).
- Sayers, J. R. & Eckstein, F. *J. biol. Chem.* **265**, 18311–18317 (1990).
- Ceska, T. A., Sayers, J. R., Eckstein, F. & Suck, D. *J. molec. Biol.* **233**, 179–182 (1993).
- Sayers, J. R. & Eckstein, F. *Nucleic Acids Res.* **19**, 4127–4132 (1991).
- Kim, Y. et al. *Nature* **376**, 612–616 (1995).
- Sayers, J. R. *J. theor. Biol.* **170**, 415–421 (1994).
- Skinner, M. M. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 2071–2075 (1994).
- Beese, L. S. & Steitz, T. A. *EMBO J.* **10**, 25–33 (1991).
- Lima, C. D., Wang, J. C. & Mondragon, A. *Nature* **367**, 138–146 (1994).
- Berger, J. M., Gamblin, S. J., Harrison, S. C. & Wang, J. C. *Nature* **379**, 225–233 (1996).
- Wigley, D. B., Davies, G. J., Dodson, E. J., Maxwell, A. & Dodson, G. *Nature* **351**, 624–629 (1991).

- Robins, P., Pappin, D. J. C., Wood, R. D. & Lindahl, T. *J. biol. Chem.* **269**, 28535–28538 (1994).
- Murante, R. S., Rust, L. & Bambara, R. A. *J. biol. Chem.* **270**, 30377–30383 (1995).
- Sayers, J. R., Krekel, C. & Eckstein, F. *BioTechniques* **13**, 592–596 (1992).
- Otwinowski, Z. DENZO (Yale University, New Haven, CT, 1993).
- Furey, W. & Swaminathan, S. *Am. Crystallographic Ass. Meet. Abstr.* **18**, 73 (1990).
- Vellieux, F. M. D. A., Hunt, J. F., Roy, S. & Read, R. J. *J. appl. Crystallogr.* **28**, 347–351 (1995).
- Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. *Acta crystallogr. A* **47**, 110–119 (1991).
- Collaborative Computational Project Number 4 *Acta crystallogr. D* **50**, 760–763 (1993).
- Brunker, A. T. X-PLOR, Version 3.1 (Yale University, New Haven, CT, 1993).
- Nicholls, A., Sharp, K. A. & Honig, B. *Proteins Struct. Funct. Genet.* **11**, 282–296 (1991).
- Carson, M. *J. molec. Graphics* **5**, 103–106 (1987).

ACKNOWLEDGEMENTS. We thank A. Thompson and A. Gonzalez for helping to collect multi-wavelength anomalous dispersion data.

CORRESPONDENCE and requests for materials should be addressed to T.A.C. (e-mail: ceska@embl.heidelberg.de).

CORRECTION

Two distinct mechanisms for long-range patterning by Decapentaplegic in the *Drosophila* wing

Thomas Lecuit, William J. Brook, Medard Ng, Manuel Calleja, Henry Sun & Stephen M. Cohen

Nature **381**, 387–393 (1996)

THE correct address of M.C. is Centro de Biología Molecular, Universidad Autónoma de Madrid, Canto Blanco, 28049 Madrid, Spain. □

ERRATUM

A late Neanderthal associated with Upper Palaeolithic artefacts

Jean-Jacques Hublin, Fred Spoor, Marc Braun, Frans Zonneveld & Silvana Condemi

Nature **381**, 224–226 (1996)

A TYPOGRAPHICAL error was introduced into the last line of Table 1(c), which should read UP – NE > NE – Hs.

Correspondence and requests for materials should be addressed to J.-J.H. (e-mail: hublin@mnhn.fr) or F.S. (f.spoor@ucl.ac.uk). □

Guide to authors of contributions to *Nature*

Nature is a weekly international journal covering all the sciences. It is intended for an interdisciplinary readership, so all manuscripts should be written clearly and simply. Contributors should pay particular attention to the fact that English is not the first language of many readers. Space is limited and competition severe, so brevity is highly valued. The length limits below are specified in terms of the number of pages taken up in *Nature*. An uninterrupted page of text contains about 1,300 words, so a typical Letter to *Nature*, occupying 2½ pages, can contain about 1,500 words of text and four modest display items (figures and/or tables) with brief legends. There are no formal limits on the number of display items, but extra display material must be accompanied by commensurate reductions in the length of the text. Length limits can, however, be varied upwards or downwards at the editors' discretion.

Nature publishes the following types of contribution:

Review Articles survey recent developments in a topical area of scientific research. They should not generally occupy more than 6 pages of *Nature*.

Progress articles are relatively informal reviews of particularly topical and fast-moving fields. They do not exceed 4 pages of *Nature*.

Contributors wishing to submit a Review Article or Progress should first send a one-page synopsis to the Reviews Coordinator, from whom guidelines are available on request.

Articles are reports of original scientific research whose conclusions represent a substantial advance in understanding of an important problem and are of broad general interest. They do not generally exceed 5 pages of *Nature* and have no more than 50 references. They have a 'heading' of up to 80 words, often rewritten by *Nature's* editors, which advertises the content of the paper in general terms and should not contain numbers, abbreviations or measurements unless central to the message of the paper. The introduction and summary should be contained in the first three or four paragraphs of the main text. Layout of methods, figures and legends is summarized below.

Letters to *Nature* are short reports focused on an outstanding finding whose implications are of interest to nonspecialists. They should not generally exceed 2½ pages of *Nature*. They have a bold first paragraph of not more than 180 words summarizing the background, rationale, conclusions and main implications of the work, aimed particularly at nonspecialist readers. Letters should have no more than 30 references. Layout of methods, figures and legends is summarized below.

Preparation of Articles and Letters

All manuscripts should be written in English, typed double-spaced on one side of the paper only. Manuscripts submitted in electronic form cannot be considered.

Manuscripts should be accompanied by a brief covering letter from the corresponding author, containing a full postal address, telephone and fax numbers, and e-mail address.

Five copies of manuscripts and original figures are needed, together with two copies of the covering letter. Five copies of manuscripts by any of the authors that have been submitted or are in press elsewhere should also be included, clearly marked as such; failure to do so may prejudice consideration of the manuscript.

Unless otherwise instructed by the editor handling a manuscript, when sending revised or resubmitted manuscripts five copies are required, each accompanied by a copy of the authors' response to referees' and editors' comments on the earlier version.

Titles should not exceed three lines of 31 characters each (including spaces), and should avoid numerical values, acronyms, abbreviations and punctuation wherever possible.

References should be numbered sequentially as they appear in the text, tables and figure legends. Only papers that have been published or submitted to a journal should be in the numbered list: manuscripts in preparation should be mentioned in the text with a list of authors. Reference lists should not contain textual material, grant details or acknowledgements.

Acknowledgements must be brief and appear after the reference list. *Nature* does not publish grant or contribution numbers.

Figures should not be larger than 22 by 28 cm unless unavoidable. They should not contain more than one panel unless the assembled parts are logically connected. Each copy should be marked with the corresponding author's name and, when known, the manuscript reference number. One photocopy of the original figures should be provided. All original figures are returned when *Nature* cannot offer to publish a manuscript, but one copy of the manuscript and a

photocopy of the figures are kept in confidential files for three months and then destroyed. Detailed instructions for figure preparation are available from the production department, including details of acceptable digital formats.

Colour figures are welcome but a contribution towards the cost of reproduction is requested. Inability to pay this charge will not prevent *Nature* publishing figures where colour is essential.

Tables. Symbols and abbreviations in a table should be defined immediately below the table, followed by essential descriptive material, all in double-spaced text. Tables should each be presented on a separate sheet of paper.

Methods. Brief details of materials and methods can be given in the text. Otherwise, they should be described at the end of the text in a section entitled "Methods", subdivided by bold headings referring to the methods used. This section should not exceed 800 words and ideally should be shorter. Additional material can be deposited on *Nature's* Web site. In contributions with methods sections, figure legends should not exceed 150 words each, and should be confined to a description of the figure and the symbols used.

Materials. As a condition of publication, authors are required to make materials and methods used freely available to academic researchers for their own use. Supporting datasets must be made available at the time of publication from the authors directly, by posting on *Nature's* web site, by deposition in the appropriate public database or by distribution on the Internet. In the case of X-ray crystallographic coordinates only, public access may be delayed for up to one year after publication.

Articles and Letters should both contain a statement at the end of the text: "Correspondence and requests for materials to xxxxx", with an e-mail address if desired, followed by the accession numbers or Web site address for supporting datasets. The designated author need not necessarily be the senior author, but *Nature* will expect the named author to handle matters arising from published contributions.

Manuscripts can be submitted to the Editor at *Nature*, Porter's South, Crinan St, London N1 9XW or at *Nature*, 968 National Press Building, Washington, DC 20045-1938. Proofs should be returned by express mail to London. Fax numbers and e-mail addresses for all editorial offices are on *Nature's* masthead.

Other contributions

Nature publishes informal contributed material as follows:

Correspondence contributions are short comments on topical issues, anecdotal material or reactions of readers to nontechnical material published in *Nature* (mainly News, Briefings, Opinion, Book Review or Commentary articles). All are unsolicited; submissions by e-mail to corres@nature.com are preferred (accents on names and places should be clearly indicated).

Commentary articles are informal, editorial comments on topical issues of public interest that have some bearing on scientific research. They are often controversial. Unsolicited manuscripts are considered, but enquiries can be made to the Commentary Editor by e-mail (nature@nature.com) or fax as a one-page synopsis or informal letter before formal submission.

News and Views articles inform nonspecialist readers about new scientific advances, either as reported in published papers or in the form of meeting reports. Most are commissioned, but proposals can be made to the News and Views Editor in advance of publication of the paper or of the meeting concerned. Authors are not allowed to discuss their own work or work from elsewhere in their institution. Detailed guidelines are available on request.

Scientific Correspondence is a predominantly peer-reviewed section of the journal specializing in very brief reports of general interest and discussion of topical issues. Priority is given to contributions of fewer than 500 words. Detailed guidelines are available on request. All contributions are unsolicited, and should be submitted by mail, not e-mail. Copies of contributions submitted by fax should be sent by mail with original figures and a cover letter noting that the manuscript has been faxed.

Book Reviews are all commissioned by the Book Review Editor. Unsolicited contributions are not considered.

Nature's editors will request final copy for Commentary, News and Views, Scientific Correspondence and Book Reviews on disk or via e-mail, as appropriate.

A more detailed Guide to Authors is available on *Nature's* WWW sites at <http://www.nature.com> or <http://www.america.nature.com>

Biochemical exchange

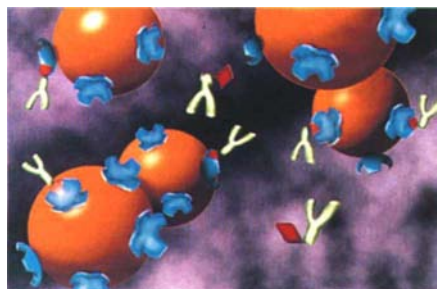
This edition highlights a new protein that enhances cell proliferative properties, a selection of protein kinases, active peptides conjugated to gold particles, a test for salivary melatonin levels and a cyanide-free lysing agent.

WITA provides a selection of **biological services** in two-dimensional gel electrophoresis, protein sequencing and peptide synthesis. Up to 10,000 protein species from complex protein mixtures, such as tissues or cells, may be separated from as little as 2 mg of starting material. Proteins in the low picomole range can be sequenced, analysed by mass spectrometry and identified by comparisons with sequence databases. Additionally, biochemical methods and mass spectrometry should help with the elucidation of protein function, as well as the detection of post-translational modifications of the investigated protein. For immunological studies antibodies are produced against the peptides, which are synthesized by an automated peptide synthesizer. This kind of analysis can be used for the comparison of two or more biological situations, for example, different differentiation steps in diseased tissue compared with healthy tissues, or the comparison of cells with and without treatment of pharmaceuticals at different phylogenetic stages.

Reader Enquiry No. 100

Peptides and proteins

Bangs Labs has announced the availability of ProActive **protein-coated microspheres**, including streptavidin, avidin, monomeric avidin, protein A and goat anti-mouse coatings on superparamagnetic and non-magnetic polystyrene ('latex' particles) or silica microspheres. According to the manufacturer, these protein coatings



ProActive protein A-coated microspheres.

enable easy coupling of antigens, antibodies, haptens or oligonucleotides for a wide variety of immunoassays. The company states that users can quickly couple biotinylated ligands to the streptavidin beads or attach mouse monoclonal antibodies to the goat anti-mouse particles. Polyclonal anti-

bodies can be oriented properly by binding to protein A-coated microspheres. Custom orders are encouraged for any appropriate protein on polystyrene, silica or magnetic microspheres from ~20 nm to ~1 mm in diameter.

Reader Enquiry No. 101

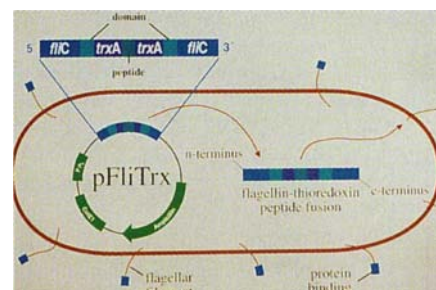
B/T SciTech now offers a **GCR 1003 protein** that has been discovered to have cell proliferative properties similar to FBS. GCR 1003, which has been isolated from several species of plant, has been shown to sustain cell growth and proliferation and can be used in various cell culture applications. It is being sold in combination with amino acids as Prolifix and can be used in combination with cell culture media either with or without FBS. With FBS, it is added to a final concentration of 5–10 per cent. Used as a supplement, this protein is said to provide better cell growth than is achieved with FBS alone. Without FBS, it can be added to cell culture media alone, giving similar growth characteristics to FBS (± 10 per cent) and creating a completely defined medium, free of human or animal protein, but sustaining normal cell growth rates.

Reader Enquiry No. 102

TCS Biologicals now offers **Janus protein kinases (Jaks)**, which were initially identified in attempts to clone novel protein tyrosine kinases. To date, four Jaks have been identified, including the ubiquitously expressed Jak1, Jak2, Tyk2 and Jak3, the latter of which is primarily expressed in haematopoietic cells. Their importance in cellular signalling emerged with the demonstration that they physically associate with one or more of the receptor chains, that ligand binding rapidly induces their tyrosine phosphorylation and that associated with these phosphorylations, their kinase activity is dramatically increased. The availability of highly specific reagents for the Jak proteins should help unravel their role in signalling through the cytokine receptor superfamily. Jaks may also contribute to signalling through other types of receptor systems; recent studies have implicated Jak signalling through the angiotensin II receptor.

Reader Enquiry No. 103

The **FliTrx random peptide display library** from Invitrogen is the company's latest innovation in the study of



Invitrogen's FliTrx random peptide library.

protein-protein interactions. The system utilizes bacterial flagella to display random peptide libraries on the surface of *Escherichia coli*. Unlike phage-based display methods, phage infection and isolation steps are not necessary. The random peptides are positioned in a flagellum (Fli) thioredoxin (Trx) fusion protein, which is exported and assembled into partly functional flagella. Proteins of interest are screened by panning and then plasmids isolated from positive clones are sequenced. The library (1.77×10^8 primary clones) is supplied in five 1-ml aliquots containing 2×10^{10} CFU ml⁻¹.

Reader Enquiry No. 104

Nanoprobes has announced the synthesis of active peptide-Nanogold conjugates, and now provides reagents for investigators to make a variety of **peptide conjugate probes**. These probes may be used to study active peptide binding to isolated molecules by high-resolution electron microscopy, or to visualize the peptide distribution at the cellular level using light or confocal microscopy and silver enhancement. Nanogold is a small 1.4-nm gold cluster with an organic shell and a mono-functional reactive arm, either a maleimide (which reacts with thiols) or a sulphydroxy-succinimide ester (which reacts with amines). Covalent linking to peptides may be achieved by reacting with the α -NH₂ group or, for example, an added terminal cysteine. The conjugates may be tracked easily in gels or western and other blots by developing with the company's silver reagents (LI Silver or HQ Silver).

Reader Enquiry No. 105

Heliosynthese has developed a selection of therapeutic molecules extracted from microalgae, including **superoxide dismutase containing manganese (Mn-SOD)**. According to the manufacturer, the thera-

peutic value of SOD has proven effective in the treatment of numerous oxidative stress illnesses, as an adjuvant therapy in AIDS and in the prevention of certain radiotherapy side effects. The company currently has the ability to produce industrial-scale quantities of Mn-SOD from strains of microalgae (*Porphyridium cruentum*). This process is carried out in continuous flow under sterile, autotrophic conditions using a bioreactor and processes. **Reader Enquiry No. 106**

Pharmacia Biotech now offers **histidine-tagged fusion proteins**, which can be prepared with high purity in a single, rapid step using the new HisTrap purification kit. Fusion proteins are purified directly from bacterial lysates and are recovered from the matrix under mild elution conditions, in order to preserve the antigenicity and functionality of the protein.



Pharmacia's HistoTrap purification kit.

The kit contains ready-to-use columns, accessories and buffer concentrates sufficient to perform 10–12 purifications when operated with a syringe. **Reader Enquiry No. 107**

A new ultrasensitive, direct ^{125}I RIA test for **salivary melatonin** is now available from Bühlmann Laboratories. The neuroendocrine hormone melatonin modulates circadian biorhythms and is an excellent marker for the setting of the endogenous clock. Salivary melatonin level is lower than that in serum, but is said to be closely correlated. The test allows simplified and convenient diagnosis of shifted melatonin profiles in conditions such as seasonal affective disorders, delayed sleep-phase syndrome, insomnia or depression related to disturbed circadian rhythms. The kit combines convenient sampling with high sensitivity (0.2 pg ml^{-1}) and a standard range of $0.5\text{--}50 \text{ pg ml}^{-1}$. **Reader Enquiry No. 108**

Hycor Biomedical now offers Hytec, an automated EIA system for **autoimmune testing in an ELISA format**. The system performs anti-RF, anti-dsDNA and ENA screening for six antigens, as well as confirmatory tests for anti-SS-A, anti-SS-B, anti-Sm, anti-Sm/RNP, anti-Scl-70 and anti-Jo-1. Assays for autoantibodies to cardiolipin and antigens associated with thyroid, liver and other autoimmune disorders are under development. Manual and semi-automated kit configurations are also available. The assays are stated to demonstrate

improved performance with correlation to western blotting for positive samples of >99 per cent, with coefficients of variation typically less than nine per cent. The manufacturer states that once the samples and reagents are loaded, the instrument performs all sample dilution and dispensing, reagent addition, incubation, washing, reading and results calculation without operator intervention. Autoimmune results for one run of up to 82 patients and six antigens are reported in 3.5 hours. The ENA Screen and confirmatory tests may be run simultaneously. **Reader Enquiry No. 109**

Kamiya Biomedical is now offering several new **monoclonal antibodies for multidrug resistance (MDR)** research. These include anti-human multidrug resistance-associated protein (MRP) mAbs (clones MRPr1 and MRPM6), anti-human LRP (multidrug resistance-related antigen) mAb (clone LAP-56) and anti-human topoisomerase II α mAb (clone 4/2D), which can be used for studying non-P-glycoprotein MDR. Applications include immunohistochemistry, flow cytometry and western blotting. For studying P-glycoprotein MDR, the company also offers anti-human P-glycoprotein mAb (clone MRK16) and anti-human P-glycoprotein (clone 6/1 C). Here, applications include flow cytometry, western blotting, immunohistochemistry and immunoprecipitation. These antibodies are currently available for research use only. **Reader Enquiry No. 110**

Among Mallinckrodt Baker's range of **reagents for cell analysis** is a cyanide-free lysing reagent, Cymet MEK III Diff. Supplied in one-litre containers, the reagent is designed for use with Nihon MEK 6108 series cell analysers. It contains phenoxypolyethoxy-ethanol rather than sodium dodecyl sulphate and/or KCN, and is considered non-hazardous, causing no damage to surface water or sewer systems. This product should be used only in combination with Diluid III Diff, which is available from the company in 20-litre polycubes. A full evaluation report, available separately on request, also indicates that the use of enzymatic cleaning solution retards deterioration of instrument tubing and other sensitive components. **Reader Enquiry No. 111**

Research Biochemicals has introduced a new series of products specifically targeted for the generation of short-lived **radio-labelled tracers for positron emission tomography (PET)**. The line-up includes reagents and reference standards for the most commonly used PET radiotracers, including [^{18}F]-2-fluoro-2-deoxy-glucose and [^{18}F]-6-fluoro-DOPA. These products

are stated to be of high purity, following extensive quality control to assure lot-to-lot consistency. To simplify setup and preparation time, each reagent is supplied in pre-measured, single-use packaging, thus eliminating waste and the risk of contamination that may be present with large stock bottles. All precursors have expiration dates to ensure consistent quality and to comply with regulatory requirements. Formulations specific for computer-controlled synthesis devices, such as those supplied by Siemens, General Electric and Ion Beam Applications, are also available. **Reader Enquiry No. 112** □

These notes are compiled by Brendan Horton from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal.

ADVERTISEMENTS

BIANORM®
EQUILIBRIUM DIALYSIS SYSTEM

The gentlest method to quantify the binding parameters and thermodynamic data between ligands and biopolymers. Results come closest to the actual values. Sample volume from 0.2 ml to 5.0 ml. Up to 20 experiments in parallel.

BIANORM
GIVE YOUR DIALYSIS A BIANORM ADVANTAGE

READER ENQUIRY NO 15

Menar Organics Limited
Unit 5, Menai Technology Centre, Deiniol Road, Bangor, Gwynedd, N. Wales LL57 2UP.
Telephone (01248) 354103 Ext 205

RARE ORGANIC CHEMICALS FOR LABORATORY USE AND CUSTOM SYNTHESIS

Contact us now for your free copy of our catalogue of screening compounds, also available on 3.5" disk.

READER ENQUIRY NO 16

DNA SEQUENCING

£45 per reaction

- PCR products
- M13 inserts
- plasmid inserts
- single strand DNA
- double strand DNA
- no result? no charge!

By research labs for research labs

Molecular Biology Services
King's College School of Medicine and Dentistry
Rayne Institute, 123 Coldharbour Lane, London SE5 8NU
Tel: 0171-346 3126 Fax: 0171-733 3877

READER ENQUIRY NO 11